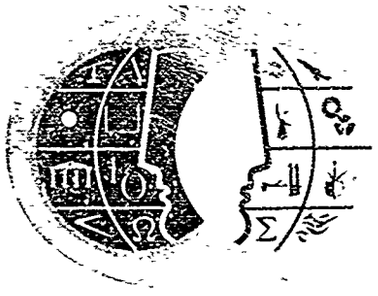


Report No. 70-32
April 1970

ON THE IDENTIFICATION OF INERTIAL MEASUREMENT UNIT ERROR PARAMETERS

J. Baziw

Report No. 70-32
April 1970

ON THE IDENTIFICATION OF INERTIAL MEASUREMENT UNIT ERROR PARAMETERS

by

John Baziw

School of Engineering and Applied Science
University of California
Los Angeles, California

PREFACE

With the development of advanced Air Force Ballistic and Space Systems there arises the requirement for increasingly sophisticated guidance and control techniques. Fundamental to this area is the identification of inertial measurement unit error parameters. This report is one of the most comprehensive documents on the subject prepared to date presenting numerous significant new results of direct interest to Air Force systems in the ballistic missile, space, and tactical systems area. Numerous fundamental results are developed and presented, virtually all of which are confirmed by extensive computational studies presented herein.

The research described in this report "On the Identification of Inertial Measurement Unit Error Parameters," No. 70-32, by John Baziw, was carried out under the direction of C. T. Leondes, E. B. Stear, A. R. Stubberud, and D. A. Wismer, Co-Principal Investigators in the Schools of Engineering in the University of California at Los Angeles, Irvine, and Santa Barbara.

This research project is supported by the Air Force Space and Missile Systems Organization under Contract No. F04701-69-C-0182, Advanced

Missile and Space Systems Studies, and the Air Force Office of Scientific Research under Contract No. AFOSR 699-67, Basic Control Systems.

This report was the basis for a Doctor of Philosophy dissertation submitted by the author.

ABSTRACT

The problem of identifying the error parameters associated with inertial measurement units is considered in this report. This is an important practical problem which is included in a large class of system parameter identification problems. A general approach for formulating the many possible inertial measurement unit (IMU) error parameter configurations is given, and specific realizations are specified in detail. The formulation is such that time correlated environmental and observational random disturbances can be incorporated. Experimental results showing the effects of state and observation noise power levels, and the nominal trajectory on the identification of the error parameters for three specific configurations are presented. These results indicate that a meaningful optimization problem can be formulated in terms of the nominal trajectory variables. The problem is then considered as an optimal control problem with the cost being a functional of the estimation covariance matrix and the controls, where certain nominal trajectory variables are considered as the controls. The question of the existence of optimal controls, the necessary conditions which the optimal controls must satisfy, and the computational aspects for computing the optimal controls are considered.

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

FIGURES

FIGURES

FIGURES

NOMENCLATURE

E_n n-dimensional Euclidean space

$x, \underline{x}$ n-dimensional vectors

$A, [A]$ n x m-dimensional matrices

$\underline{x}^*, A^*$ transpose of a vector and a matrix, respectively

$\langle x, y \rangle, \|x\|_2^2$ inner product of two vectors

t, t_o, T, I time, initial time, final time, and the time interval $[t_o, T]$, respectively

$\Phi(t, \tau)$ fundamental solution (transition matrix) of a linear differential equation

$IMU_{(\cdot)}$ inertial measurement unit $(\cdot)$

$B^{(j)}$ object j in the system, with j = m, o, s, r, and i, respectively, to denote the vehicle, a reference IMU (IMU_O), a slaved IMU (IMU_S), a, generally, non-stationary reference point, and an inertially fixed reference point, respectively

$B.P.^{(j)}$ a fixed point of object j

$\underline{X}_\alpha^{(j)}, \alpha = 1, 2, 3,$ orthogonal axes centered at $B.P.^{(j)}$

$\underline{b}^{(j,k)}, \epsilon\underline{b}^{(j,k)}$ vector distance from $B.P.^{(j)}$ to $B.P.^{(k)}$, and the error in the knowledge of $\underline{b}^{(j,k)}$, respectively

$\underline{\omega}, [\omega]$ angular velocity vector and matrix, respectively

$\underline{a}, [a]$ acceleration vector and matrix, respectively

$[\Omega], [\Omega']$ especially defined angular velocity matrices

$\theta_\alpha^{(j,k)}$ $\alpha = 1, 2, 3,$ body angles from $\underline{X}_\beta^{(j)}$ to $\underline{X}_\beta^{(k)}$

$T^{(j,k)}$ orthogonal coordinate transformation from $B^{(j)}$ to $B^{(k)}$

$\underline{n}$ random position changes

$\mathring{\underline{x}}$ a time derivative with respect to a non-stationary coordinate system

$\Delta \underline{a}_s^{(j,k)}$ difference in sensed acceleration between $B^{(j)}$ and $B^{(k)}$

$\Delta \underline{a}_m^{(j,k)}$ difference in measured acceleration between $B^{(j)}$ and $B^{(k)}$

$\underline{v}_m^{(o,s)}$, $\underline{p}_m^{(o,s)}$ first and second integrals, respectively, of the measured acceleration between IMU_s and IMU_o

$\underline{\eta}, \underline{n}, \underline{\nu}, \underline{\nu}_c, \underline{\nu}_w$ various random processes

$\underline{\Psi}, \underline{\Psi}(t_o)$, Ψ_i, $i = 1, 2, 3$ vector of initial misalignment angles

$\underline{K}_a$ vector of the (six) accelerometer error parameters

$\underline{K}_g$ vector of the (six) gyroscope error parameters

$\underline{K}^m$ vector of various measurement error parameters

$\underline{k}_S$, k_{S_i}, $i = 1, 2, 3$, mass unbalance errors along the spin axis for three gyros

$\underline{k}_I$, k_{I_i}, $i = 1, 2, 3$, mass unbalance errors along the input axis for three gyros

$\underline{\epsilon}$, ϵ_i, $i = 1, 2, 3$, constant drifts for three gyros

$S_x(\omega), R_x(\tau)$ power spectrum and autocorrelation, respectively, for the random process x

Configuration IA Identification of $\underline{\Psi}$, with correlated random acceleration

Configuration IB Identification of $\underline{\Psi}$, with uncorrelated random acceleration

Configuration II Identification of $\underline{\Psi}$ and $\underline{\epsilon}$

Configuration III Identification of $\underline{\Psi}$, and the mass-unbalance gyro drifts, $\underline{k}_{m.u.}$

$\sigma_n^2 = \phi_o$, the correlated random acceleration power

$\sigma_R^2 \cdot I$ the uncorrelated random acceleration covariance matrix

$\sigma_W^2 \cdot I$ the observation noise covariance matrix.

$\sigma_{\hat{\psi}_i}$ the standard deviation in the estimated misalignment angles

$\sigma_{\hat{\epsilon}_i}$ the standard deviation in the estimated constant gyro drift rates

$\sigma_{\hat{k}_{S_i}}, \sigma_{\hat{k}_{I_i}}$ the standard deviation in the estimated mass-unbalance gyro drift terms

$E[\underline{x}]$, $\mathrm{Cov}(\underline{x}, \underline{y})$ expected value of $\underline{x}$, and covariance of $\underline{x}$ and $\underline{y}$, respectively

$P = E[(\underline{x} - \hat{\underline{x}})(\underline{x} - \hat{\underline{x}})^*]$ covariance matrix of errors in the estimated value $\hat{\underline{x}}$ from the true value $\underline{x}$

p_{ij} elements of the covariance matrix P

P^{xy} covariance matrix associated with the x and y portions of a vector z, where $z = (x, y)^*$

$\|P\|_i$, $i = 1, 2, 3, \infty$, various norms in E_{n^2} associated with matrix P

R, Q covariance matrices of the time uncorrelated observation and state vector noise disturbances.

$\underline{u}, \mu$ an m-vector of control functions and a scalar-valued control function, respectively

M_i various constraints on the control function $\underline{u}$

$U(t) \subset E_m$ a pointwise constraint set for $\underline{u}(t)$ for $t \in I$

$\mathcal{U}$ a set of control functions

J a cost functional associated with P, $\underline{u}$, and T

$\mathrm{Trace}\left[WP(T; \underline{u})\right]$ sum of the diagonal elements of WP evaluated at $t = T$, with the control $\underline{u}(t)$, $t_o \leq t \leq T$

$\{\underline{u}_k\}$ a sequence of functions

$\underline{u}_k \xrightarrow{W} \bar{\underline{u}}$ weak convergence of $\{\underline{u}_k\}$ to $\bar{\underline{u}}$

$L_2([a, b])$ space of square integrable vector functions on the time interval [a, b]

H Hamiltonian function associated with P $\underline{u}$ and J

Λ, λ_{ij}, vector of adjoint variables associated with the covariance matrix P, and the elements of Λ, respectively

$I(\Lambda_0)$, $I(P_T)$ cost functional associated with the error in guessing the initial adjoint matrix $\Lambda(t_0) = \Lambda_0$, and in guessing the final covariance matrix $P(T) = P_T$, respectively

W a non-negative weighting matrix associated with P.

CHAPTER I

INTRODUCTION

The problem of identifying the error parameters associated with inertial measurement units (IMU) is considered in detail. This is an important practical problem which is included in the general class of problems formulated in the next section. A general approach for formulating the many possible IMU error parameter configurations is given in Chapter II, and specific realizations are given in Chapter III. Minimum variance estimation techniques for the error analysis of the identification procedure, and typical experimental results are presented in Chapters IV and V. In Chapter VI, techniques from optimal control theory are used to characterize the best nominal trajectory from a class of admissible nominal trajectories. This trajectory is best in the sense that the identification of the error parameters is the most satisfactory under given physical conditions. Computational aspects for obtaining the optimal nominal trajectory are discussed in Chapter VII.

1.1 The General Problem

Many physical systems might be described by the state equation

$$\dot{y} = f(t, y, u_1, w_1, \alpha_1) \tag{1.1}$$

and the observation equation

$$z = h(t, y, u_2, w_2, \alpha_2) + v \tag{1.2}$$

In these equations all the variables are generally vectors, except for the time t. The n-dimensional vector function f and the m-dimensional vector function h are assumed to be sufficiently smooth so that the required linearizations which are subsequently required are valid. The specific definitions are as follows:

y is an $n \times 1$ vector of state variables,

u_1 is an $r_1 \times 1$ vector of state control variables,

w_1 is a $p_1 \times 1$ vector of random state disturbances,

α_1 is a $k_1 \times 1$ vector of constant state parameters,

z is an $m \times 1$ vector of observation variables,

u_2 is an $r_2 \times 1$ vector of observation control variables,

w_2 is a $p_2 \times 1$ vector of random observation disturbances,

α_2 is a $k_2 \times 1$ vector of constant state parameters,

v is an $m \times 1$ vector of random white noise disturbances.

The vectors α_1 and α_2 of system parameters are assumed to be made up of constants, although smooth functions would also be considered in this framework by making suitable polynomial approximations, and then increasing the dimension of α_1 and α_2 appropriately. The random disturbances w_1 and w_2 may be time-correlated or not, but it is assumed that the statistics of these vectors are completely known and that they may be modelled by white noise passing through a linear dynamical system (Markovian property).

The general problem which might be presented is to identify the system parameters α_1 and α_2 based on the observations z. Of course, it is also usually of interest to obtain estimates of the state y. The controls u_1 and u_2 are functions which are available to change the evolution of the state and observation equations. This change in these equations is to be made in such a way that the identification might be accomplished in some optimal fashion, subject to certain constraints in the controlling functions u_1 and u_2. Except for special theoretical questions, there is not much that can be said about the analytical properties of this problem in this generality. Certain linearizations might be made to make the problem more tractable and perhaps amenable to certain mean square optimization techniques.

Under suitable assumptions, the system (1.1) , (1.2) may be written as

$$\dot{y} = A^{y}(t,u_1)\, y + B^{w_1}(t,u_1)\, w_1 + C^{\alpha_1}(t,u_1)\alpha_1 \qquad (1.3)$$

and

$$z = M^{y}(t,u_2)\, y + B^{w_2}(t,u_2)\, w_2 + C^{\alpha_2}(t,u_2)\alpha_2 + v \qquad (1.4)$$

These equations represent small deviations about nominal or expected values of the variables described above, which are denoted by $\bar{y}, \bar{w}$, etc. The matrices in Equations (1.3) and (1.4) are the various partial derivatives of the functions f and h, and are evaluated about the nominal or expected values. Typically,

$$A^{y}(t,\, u_1) \equiv \left.\frac{\partial f(t,\, \cdot\,)}{\partial y}\right|_{\bar{y},\, \bar{w}_1, \bar{\alpha}_1}$$

and

$$B^{w_1}(t,\, u_1) \equiv \left.\frac{\partial f(t,\, \cdot\,)}{\partial w_1}\right|_{\bar{y},\, \bar{w}_1, \bar{\alpha}_1}$$

It is noticed that we do not linearize about the controls u_1 and u_2 because in the application which is considered here, the nominal values $\bar{u}_1$, $\bar{u}_2$ are also to be chosen in an optimal way. It is next assumed that w_1, w_2 have the Markovian representations $\dot{w}_1 = A^{w_1} w_1 + B^{w_1'} v_1$ and $\dot{w}_2 = A^{w_2} w_2 + B^{w_2'} v_2$.

If we write $\dot{\alpha}_1 = 0$ and $\dot{\alpha}_2 = 0$, then the above equations might be collected together and written as

$$\dot{Y} = AY + Bw \tag{1.5}$$

and

$$z = MY + v \tag{1.6}$$

where

$$Y \equiv (y, w_1, w_2, \alpha_1, \alpha_2)^*$$

$$A = A(t, u_1) = \begin{bmatrix} A^y & B^{w_1} & 0 & C^{\alpha_1} & 0 \\ 0 & A^{w_1} & 0 & 0 & 0 \\ 0 & 0 & A^{w_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$w = (w_1, w_2)^* \qquad B = \begin{bmatrix} 0 & 0 \\ B^{w_1'} & 0 \\ 0 & B^{w_2'} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and

$$M = M(t, u_2) \equiv \begin{bmatrix} I & 0 & B^{w_2} & 0 & C^{\alpha_2} \end{bmatrix}$$

If the controls u_1, u_2 have been prespecified, the problem formulated by equations (1.5) and (1.6) is in the standard form for applying the well-known minimum variance estimation techniques [1]. Assuming the random vectors possess the known statistical properties

$$E[w(t)] = E[w(\tau)] = 0, \quad E[w(t)\, v(\tau)^*] = 0,$$

$$E[w(t)\, w(\tau)^*] = Q(t)\delta(t-\tau), \quad E[v(t)\, v(\tau)^*] = R(t)\delta(t-\tau),$$

$$E[Y(t_o)] = 0, \quad E[Y(t_o)\, Y(t_o)^*] = P_o,$$

$$E[Y(t_o)\, v(\tau)^*] = 0, \quad E[Y(t_o)\, w(\tau)^*] = 0,$$

then the minimum variance estimate $\hat{Y}$ is given by the solution of the differential equation

$$\frac{d\hat{Y}}{dt} = \left[A - KM\right]\hat{Y} + Kz \; ; \quad Y(t_o) = 0 \qquad (1.7)$$

The optimal gain matrix K is given by

$$K = PM^* R^{-1}$$

where the covariance matrix P, which is defined by

$$\text{Trace } P = \text{Trace } E\left[(Y - \hat{Y})(Y - \hat{Y})^*\right],$$

satisfies the matrix Riccati equation

$$dP/dt = AP + PA^* - PM^* R^{-1} MP + BQB^* ; \qquad (1.8)$$

$$P(t_0) = E[Y(t_0) Y(t_0)^*] = P_0$$

The dependence of the solution on the control parameters $u \equiv (u_1, u_2)$ will be denoted by $\hat{Y}_u$ and P_u. The optimal control problem which naturally arizes is to minimize some functional of P, for example $\min_u \text{trace}[WP(T;u)]$ where W is a weighting matrix and T is the end time, or else to choose a u such that certain diagonal elements of P fall into a prespecified region in minimum time.

For a linearized minimum variance estimator this cost would depend on a particular stochastic realization of the system. The variances in the resulting P(T) for particular realizations would then need to be considered to decide on a suitable "average" cost. Thus a linearized minimum variance estimator would complicate the optimization considerably. Since we should like to apply some of the known techniques from optimal control theory [2], it shall be assumed that we have a linear system from the outset. This is indeed the case for the IMU error identification problem, because the design of the associated instruments is such that it keeps all errors reasonably small. So that even though the various errors might be very complicated in nature, their contribution to the system dynamics can be accurately represented by a linear transformation.

1.2 The IMU Error Parameter Identification Problem

The problem of identifying the error parameters associated with inertial measurement units (IMU's) arizes in many applications associated with the inertial navigation of ships, airplanes, missiles, and space vehicles. Investigation of this specific class of important problems provides motivation and insight into the general problem outlined above. A discussion of the error parameters which are typically associated with IMU's and the physical environment in which such systems operate is given in Chapter II. A general formulation of the problem and a discussion of experimental results is given in Chapters II to V.

Generally, the IMU error identification problem has the form

$$\begin{aligned} \dot{x} &= A^x x + B^n n + C^\alpha \alpha + w^x \\ \dot{n} &= A^n n + w^n \\ \dot{\alpha} &= 0 \\ z &= x + v \end{aligned} \tag{1.9}$$

Equations (1.7) and (1.8) then become, respectively,

$$\begin{aligned} d\hat{x}/dt &= (A^x - P^{xx}R^{-1})\,\hat{x} + B^n \hat{x} + C^\alpha \hat{\alpha} + P^{xx}R^{-1} z \\ d\hat{n}/dt &= -P^{nx}R^{-1}\,\hat{x} + A^n \hat{n} + P^{nx}R^{-1} z \\ d\hat{\alpha}/dt &= -P^{\alpha x}R^{-1}\,\hat{x} + P^{\alpha x}R^{-1} z \end{aligned} \tag{1.10}$$

and

$$\frac{d}{dt} P^{xx} = A^x P^{xx} + P^{xx} A^{x*} + B^n P^{nx} + P^{xn} B^{n*} + C^\alpha P^{\alpha x} + P^{x\alpha} C^{\alpha*} - P^{xx} R^{-1} P^{xx} + Q^{xx}$$

$$\frac{d}{dt} P^{nn} = A^n P^{nn} + P^{nn} A^n - P^{nx} R^{-1} P^{xn} + Q^{nn}$$

$$\frac{d}{dt} P^{\alpha\alpha} = -P^{\alpha x} R^{-1} P^{x\alpha} \qquad (1.11)$$

$$\frac{d}{dt} P^{xn} = A^x P^{xn} + P^{xn} A^{n*} + B^n P^{nn} + C^\alpha P^{\alpha n} - P^{xx} R^{-1} P^{xn}$$

$$\frac{d}{dt} P^{x\alpha} = A^x P^{x\alpha} + B^n P^{n\alpha} + C^\alpha P^{\alpha\alpha} - P^{xx} R^{-1} P^{x\alpha}$$

$$\frac{d}{dt} P^{n\alpha} = A^n P^{n\alpha} - P^{nx} R^{-1} P^{x\alpha}$$

where the covariance matrix P has been partitioned in the obvious way

$$P = \begin{bmatrix} P^{xx} & P^{xn} & P^{x\alpha} \\ P^{nx} & P^{nn} & P^{n\alpha} \\ P^{\alpha x} & P^{\alpha n} & P^{\alpha\alpha} \end{bmatrix}$$

Generally, the control parameters would appear in the matrices A^x, B^n, C^α, and Q^{xx}. Thus it is difficult to study the optimization problem in detail at this level of generality, although the optimization problem can be clearly stated. If the assumptions used for special configurations (configurations IB, II, and III of Chapter III) are invoked, the optimization starts to be tractable. Briefly, these assumptions are $n = 0$ (no correlated process noise), and

diagonal noise covariance matrices $Q^{xx} = \sigma_R^2 \cdot I$ and $R = \sigma_w^2 \cdot I$. Equation (1.10) and (1.11) reduce to

$$\begin{aligned} \frac{d}{dt}\hat{x} &= -\sigma_w^{-2} P^{xx}[\hat{x} - z] + C^{\alpha}\hat{\alpha} \\ \frac{d}{dt}\hat{\alpha} &= -\sigma_w^{-2} P^{\alpha x}[\hat{x} - z], \end{aligned} \tag{1.12}$$

and

$$\begin{aligned} \frac{d}{dt}P^{xx} &= C^{\alpha}P^{\alpha x} + P^{x\alpha}C^{\alpha *} - \sigma_w^{-2}P^{xx}P^{xx} + \sigma_R^{+2} \cdot I \\ \frac{d}{dt}P^{\alpha\alpha} &= \sigma_w^{-2}P^{\alpha x}P^{x\alpha} \\ \frac{d}{dt}P^{x\alpha} &= C^{\alpha}P^{\alpha\alpha} - \sigma_w^{-2}P^{xx}P^{x\alpha} \end{aligned} \tag{1.13}$$

respectively. The trajectory control parameters are now associated with the matrix C^{α} only. The optimization problem is then to choose these trajectory control variables so that

$$\text{Trace}\left[W^{\alpha}P^{ee}(T)\right],$$

where W^{α} is a positive definite weighting matrix, is minimized at some given end time T, or else to choose the control variables such that

$$p_{jj}(T) \leq \epsilon_j\,, \quad \epsilon_j \text{ specified, } j = 1, \cdots, n;$$

in a minimum time $T = \bar{T}$.

Discussions which are concerned with the existence of the optimal controls and the necessary conditions which the optimal controls must satisfy can be made on the basis of the problem specified by Equation (1.9). However, to gain insight from numerical results and to proceed with analytical calculations, it is more fruitful to work with the simpler models.

This study has three main objects. The first is to formulate a general method for treating the many possible IMU error parameter configurations. The second is to demonstrate, through numerical experimentation with realistic error parameter configurations, under which physical conditions it is possible to identify the error parameters and that a realistic trajectory optimization problem does exist. The third object is to formulate the optimal control problem and discuss the existence of the optimal controls, the necessary conditions which the optimal controls must satisfy, and the computational aspects for obtaining the optimal controls.

CHAPTER II

GENERAL APPROACH TO INFLIGHT IDENTIFICATION OF IMU ERROR PARAMETERS

2.1 INTRODUCTION

There are many inertially guided vehicles serving as a common carrier for a smaller inertially guided vehicle. Typically, the smaller vehicle is deployed either during the common carrier's travel or at the carrier's nominal terminus. As a contrivance to introduce the problem, consider the Saturn IV/Command and Service Module (CSM). Prior to launch of an Apollo mission it may be difficult to calibrate and/or align the CSM (slave) IMU because of its physical arrangement. The master IMU (IMU used throughout boost powered flight) is assumed to be accurately aligned and calibrated. When the vehicle is put into operation, the slave IMU could be activated prior to its being used and its output observed. The observed outputs of the slave IMU and the master IMU would then be compared and the difference in the observations used as a basis for inflight calibration and alignment of the slave IMU with respect to the master IMU.

Under ideal conditions, the readings from the two IMU's would be the same. However, there are several factors which cause the readings between the master and slave IMU's to be different. These factors include master and slave IMU error parameters (gyro drift errors, accelerometer scale factor and bias errors, platform and accelerometer misalignments), accelerometer quantization errors, computer errors, and random-induced accelerations. Assuming the variances of the master IMU error parameters are negligibly small so that the master IMU accelerometers read the true sensed acceleration, and assuming availability of appropriate probabilistic descriptions of the vehicle vibrations and of the computer and accelerometer quantization errors, which are random in nature, various techniques may be used to identify the IMU error parameters. In this chapter a general approach is formulated for identifying these IMU error parameters.

2.2 Dynamic Equations

The master IMU, denoted by IMU_C, is the reference for calibrating and aligning the slaved IMU, IMU_S. It is assumed that IMU_S is, in general, a strapped-down system, thereby insuring the necessary generality in the dynamical equations for treating either gimbaled or strapped-down IMU's. In the discussion of this section, the

notation listed below is used. Later, whenever the meaning is clear, a less complicated notation is adopted.

Referring to Figure 2.1, we let $B^{(j)}$ denote an object in the system with orthogonal axes $\underline{X}_\alpha^{(j)}$, $\alpha = 1,2,3$, centered at $B.P.^{(j)}$, a point fixed in the object. The superscript j will be m for the vehicle, o for IMU_o, s for IMU_s, r for some generally non-stationary reference point, and i for an inertially fixed reference point. $\underline{b}^{(j,k)}$ is the vector from $B.P.^{(j)}$ to $B.P.^{(k)}$, where as above j, k = m, o, s, r, and i. Letting subscript o denote a nominal, or assumed position, we define $\varepsilon\underline{b}_o^{(i,s)}$ as the error vector in the nominal position of the slaved IMU, and $\varepsilon\underline{b}_o^{(i,o)}$ as the error vector in the nominal position of the master IMU. For compactness, we shall use a dot in the superscripts to indicate situations where j can either be s or o.

The position and velocity of $IMU_{(\cdot)}$, with all vectors expressed in vehicle coordinates $\underline{X}_\alpha^{(m)}$, are given respectively by

$$\underline{b}^{(i,\cdot)} = \underline{b}_o^{(i,\cdot)} + \underline{r}^{(\cdot)}$$

$$\dot{\underline{b}}^{(i,\cdot)} = \dot{\underline{b}}_o^{(i,\cdot)} + \dot{\underline{n}}(t)^{(\cdot)} + \underline{\omega} \times \underline{r}^{(o)}$$

where the dot over a vector denotes a derivative with respect to

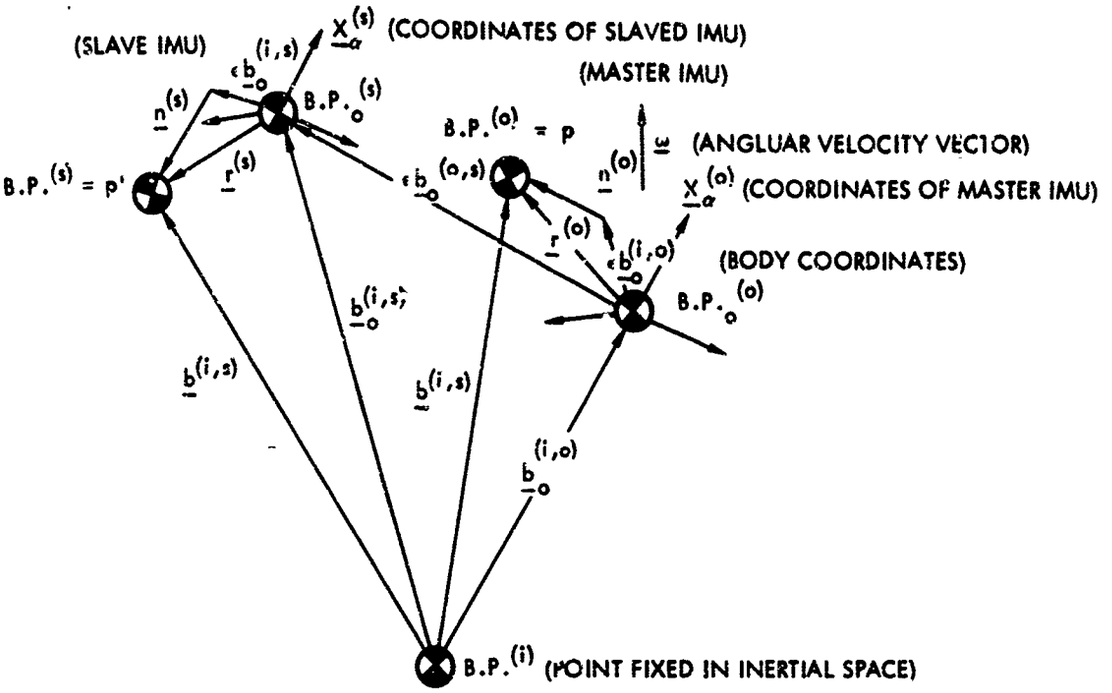


Figure 2.1 General Position of Master and Slaved IMU's

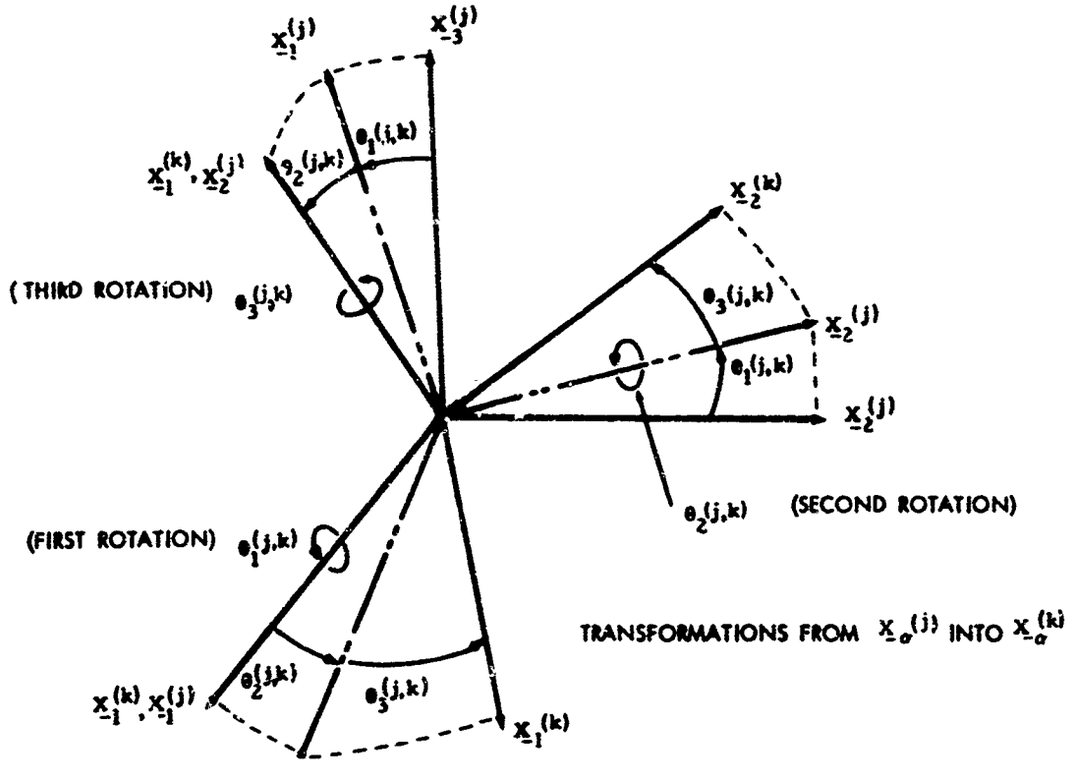


Figure 2.2 Rotations Between Coordinate Systems Using Body Angles

inertial space and the small circle over a vector denotes a derivative with respect to vehicle coordinates, $\underset{\alpha}{\overset{(m)}{x}}$. Thus the acceleration is

$$\underset{\sim}{\ddot{b}}^{(i,\cdot)} = \underset{\sim}{\ddot{b}}_o^{(i,\cdot)} + \underset{\sim}{\ddot{n}}(t)^{(\cdot)} + \underset{\sim}{\omega} x \underset{\sim}{\dot{n}}(t)^{(\cdot)} + \underset{\sim}{\omega} x \underset{\sim}{r}^{(\cdot)} + \omega x \dot{r}^{(\cdot)} + \underset{\sim}{\omega} x \underset{\sim}{\omega} x \underset{\sim}{r}^{(\cdot)}$$

so that the acceleration of the points p and p' (Figure2.1) with respect to inertial space are

$$\underset{\sim}{a}_{(\cdot)/i}^{(\cdot)} = \underset{\sim}{\ddot{b}}_o^{(i,\cdot)} + \underset{\sim}{\ddot{n}}(t)^{(\cdot)} + 2\underset{\sim}{\omega} x \underset{\sim}{\dot{n}}(t)^{(\cdot)} + (\underset{\sim}{\omega} x \underset{\sim}{\omega} x + \dot{\underset{\sim}{\omega}} x) \varepsilon \underset{\sim}{b}_o^{(i,\cdot)} + \underset{\sim}{n}(t)^{(\cdot)}$$

where the subscript $(\cdot)$ denotes either p or p'. The acceleration of IMU_S with respect IMU_O, $\Delta \underset{\sim}{a}^{(o,s)}$, is

$$\Delta \underset{\sim}{a}^{(o,s)} = \underset{\sim}{\ddot{b}}_o^{(o,s)} + \underset{\sim}{\ddot{n}}(t)^{(o,s)} + 2\underset{\sim}{\omega} x \underset{\sim}{\dot{n}}(t)^{(o,s)} + (\underset{\sim}{\omega} x \underset{\sim}{\omega} x + \dot{\underset{\sim}{\omega}} x) \underset{\sim}{n}(t)^{(o,s)} + (\underset{\sim}{\omega} x \underset{\sim}{\omega} x + \dot{\underset{\sim}{\omega}} x) \varepsilon \underset{\sim}{b}_o^{(o,s)} \qquad (2.1)$$

Equation (2.1) gives the accelerations between the master and slave IMU's due to random vehicle vibrations $\underset{\sim}{n}(t)$, and the errors in knowledge of the nominal locations of the IMU's. We can rewrite (2.1) as Equation (2.2) using the following notation:

$$[\omega] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix}$$

$$[\Omega'] = \left[[\omega]\,[\omega] + [\dot{\omega}]\right]$$

$$[\Omega] = \left[I \vdots 2\,[\omega] \vdots [\Omega']\right]$$

$$\underline{\eta}(t) = [\ddot{\underline{n}}(t), \dot{\underline{n}}(t), \underline{n}]^*$$

$$\Delta\underline{\ddot{\varepsilon}}^{(o,s)} = [\Omega]\,\underline{\eta}(t) + [\Omega']\,\varepsilon\underline{b}_o^{(o,s)} + \underline{b}_o^{(o,s)} \qquad (2.2)$$

Thus the difference in accelerations between the master and slave IMU's is composed of a term due to random vibrations of the vehicle, $[\Omega]\,\underline{\eta}(t)$; a term due to an error in knowing the true static locations of the IMU's, $[\Omega']\,\varepsilon\underline{b}_o^{(o,s)}$; and a term due to gravity and the angular velocity of the vehicle, $\ddot{\underline{b}}_o^{(o,s)}$, which is known as a function of time. Next the term $\Delta\underline{\ddot{\varepsilon}}^{(o,s)}$ is considered in the context of the measured acceleration errors.

The acceleration of $IMU_{(\cdot)}$ with respect to inertial space $\underline{a}^{(i,\cdot)}$, may be written as $\underline{a}^{(i,\cdot)} = \underline{g}^{(\cdot)} + \underline{a}_s^{(\cdot)}$ where $\underline{g}^{(\cdot)}$ is the gravitational acceleration acting on $IMU_{(\cdot)}$, and $\underline{a}_s^{(\cdot)}$ is the thrust acceleration acting at $IMU_{(\cdot)}$. Furthermore, $\underline{a}_s^{(\cdot)}$ may be written as $a_s^{(\cdot)} = a_{measured}^{(\cdot)} - \Delta\underline{a}^{(\cdot)}$ so we can write the difference in acceleration between the two IMU's as

$$\underline{a}_{measured}^{(o,s)} = -(\underline{g}^{(s)} - \underline{g}^{(o)}) - (\Delta\underline{a}_{measured}^{(s)} - \Delta\underline{a}_{measured}^{(o)})$$

$$+ [\Omega]\,\underline{\eta}(t) + [\Omega']\,\varepsilon\underline{b}_o^{(c,s)} + \ddot{\underline{b}}_o^{(o,s)}$$

Henceforth, we shall use the subscript m for the word measured. We note that the vector $\underline{a}_m^{(o,s)}$ is the quantity which may be physically observed, although perhaps indirectly, through the integral of $\underline{a}_m^{(o,s)}$. We can now write the differential equation for the measured accelerations.

The motion of the vehicle is described by the equation

$$\ddot{\underline{R}} = \underline{g}(\underline{R}, \underline{\varepsilon}_g) + \underline{f}(\underline{R}, \dot{\underline{R}}, \underline{\varepsilon}_s)$$

where $\underline{R}$, $\dot{\underline{R}}$, $\ddot{\underline{R}}$, represent the position, velocity and acceleration of the vehicle relative to an inertial reference frame, $\underline{g}(\underline{R}, \underline{\varepsilon}_g)$ is the acceleration due to gravity, $\underline{f}(\underline{R}, \dot{\underline{R}}, \underline{\varepsilon}_s)$ is the true sensed acceleration, $\underline{\varepsilon}_g$ is a vector of gravity error parameters, and $\underline{\varepsilon}_s$ is a vector of sensed acceleration error parameters.

If $\underline{R}^{(\cdot)}$ is the true position of the geometrical center of $IMU_{(\cdot)}$ then $\underline{r} = \underline{R}^{(s)} - R^{(o)}$ satisfies the differential equation

$$d^2\underline{r}/dt^2 = (\underline{g}^{(s)} - \underline{g}^{(o)}) + (\underline{f}^{(s)} - \underline{f}^{(o)}),$$

and, neglecting the effect of the error parameters $\underline{\varepsilon}_g^{(\cdot)}$,

$$d^2\underline{r}/dt^2 = (\underline{g}^{(s)} - \underline{g}^{(o)}) + \underline{a}_m^{(o,s)} + (\Delta\underline{a}_m^{(s)} - \Delta\underline{a}_m^{(o)})$$

or

$$d^2\underline{r}/dt^2 = [\Omega]\,\underline{\eta}(t) + [\Omega']\,\underline{\varepsilon b}_o^{(o,s)} + \ddot{\underline{b}}_o^{(o,s)} \qquad (2.3)$$

If $R^{(\cdot)'}$ is the computed or measured position of $IMU_{(\cdot)}$ and $\underline{r}' = \underline{R}^{(s)'} - \underline{R}^{(o)'}$, then Equation (2.3) may be rewritten in terms of $\underline{r}'$ as

$$d^2r'/dt^2 = \underline{a}_m^{(o,s)} + \underline{g}^{(o,s)'}$$

that is

$$d^2r'/dt^2 = [\Omega]\,\underline{\eta}(t) + [\Omega']\,\varepsilon\underline{b}_o^{(o,s)} - \Delta\underline{a}_m^{(o,s)} + \underline{g}^{(o,s)'} - \underline{g}^{(o,s)} \tag{2.4}$$

If $\underline{r}_o = \underline{b}_o^{(o,s)}$, the nominal or reference distance from IMU_o to IMU_s, the term $\ddot{\underline{b}}^{(o,s)}$ can be written as

$$\ddot{\underline{b}}_o = \left[\underline{g}(\underline{R}_o^{(s)}) - \underline{g}(\underline{R}_o^{(o)})\right] + \quad \underline{\omega}\times\underline{\omega}\times\underline{\omega}\times\underline{r}_o$$

It is reasonable to assume that $\underline{g}(\underline{R}^{(s)} - \underline{R}^{(o)}) \cong \underline{g}(\underline{R}_o^{(s)} - \underline{R}_o^{(o)})$, so that Equation (2.4) becomes

$$\underline{a}_m^{(o,s)} = [\Omega]\,\underline{\eta}(t) + [\Omega']\,\varepsilon\underline{r}_o + [\Omega]\,\underline{r}_o + \Delta\underline{a}_m^{(o,s)} \tag{2.5}$$

Further we can define the observable quantities $\underline{v}_m^{(o,s)}$ and $\underline{p}_m^{(o,s)}$ by $\underline{v}_m^{(o,s)} = \int\underline{a}_m^{(o,s)}$ and $\underline{p}_m^{(o,s)} = \iint\underline{a}_m^{(o,s)}$, so we can write Equation (2.5) as

$$\frac{d}{dt}\begin{bmatrix}\underline{p}_m^{(o,s)}\\ \hline \underline{v}_m^{(o,s)}\end{bmatrix} = \left[\begin{array}{c|c} 0 & I \\ \hline 0 & 0\end{array}\right]\begin{bmatrix}\underline{p}_m^{(o,s)}\\ \hline \underline{v}_m^{(o,s)}\end{bmatrix} + \begin{bmatrix}0\\ \hline [\Omega']\,\varepsilon\underline{r}_o + \Delta\underline{a}_m^{(s)} - \Delta\underline{a}_m^{(o)} + [\Omega]\underline{\eta}(t) + [\Omega]\underline{r}_o\end{bmatrix} \tag{2.6}$$

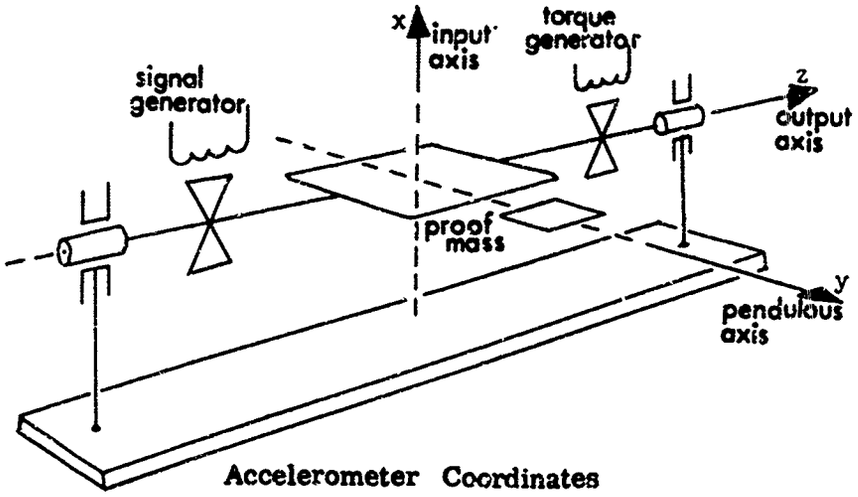
Next it is shown that the measurement errors can be written as linear combination of constants: $\Delta \underline{a}_m^{(\cdot)} = B(t)^{(\cdot)} \underline{K}^{(\cdot)}$, where $B(t)^{(\cdot)}$ is a matrix relating the error coefficients $K^{(\cdot)}$ of $IMU_{(\cdot)}$ to trajectory variables (acceleration, angular velocities, and time-varying coordinate transformations).

Let $T^{(i,\cdot)}$ denote the transformation from an inertial coordinate system to a moving (or body) coordinate system. The inverse transformation is written as $T^{(\cdot,i)}$, where the dot as before, indicates a particular object in the system. The coordinate systems specified earlier, $\underline{X}_\alpha^{(\cdot)}$, are related to one another by the transformation

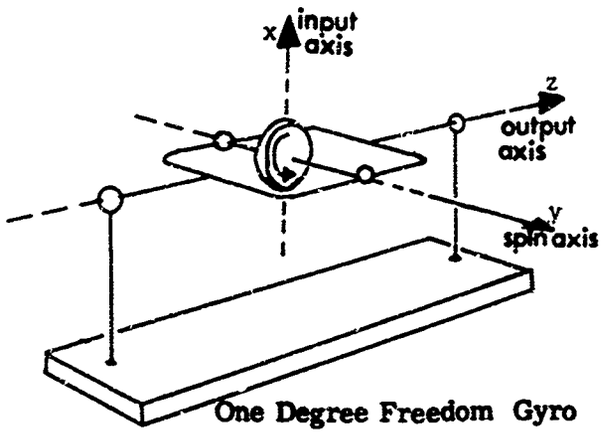
$$\underline{X}_\alpha^{(j)} = [\theta_{\alpha\beta}^{(j,k)}] \underline{X}^{(k)}, \text{ where } \alpha,\beta = 1,2,3 \text{ and } j,k = n,o,s,r,i.$$

The $\theta_{\alpha\beta}^{(j,k)}$ are elements of the usual transformation matrix, which is implicitly defined in Figure 2.2 in terms of vehicle body angles. Figure 2.3 is included to provide a set of accelerometer and gyro coordinates. In addition, two transformations, $M_o^{(o)}$ and $M_o^{(s)}$, are defined to express the transformations of the master IMU and the slave IMU, respectively, relative to some fixed inertial coordinate systems, and are constant transformations relating the initial orientations of two IMU's to some inertially fixed coordinate system.

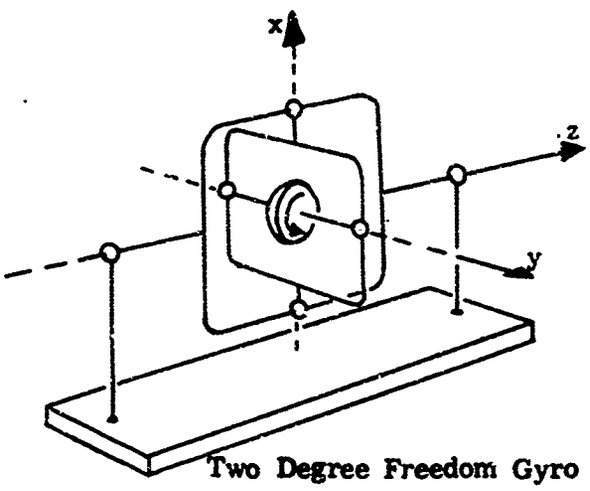
It is next necessary to define the time-varying transformation $M^{(s)}(t)$ relating the slaved IMU's present position relative to its original



Accelerometer Coordinates



One Degree Freedom Gyro



Two Degree Freedom Gyro

If a torque is applied to the inner gimbal about the x-axis, the wheel and inner gimbal will process about the z-axis. If a torque is applied to the outer gimbal about the z-axis, the wheel and inner gimbal will process about the x-axis.

Figure 2.3 Accelerometer and Gyro Coordinates

position as specified by $M_o^{(s)}$. (This is due to the assumption of a strapped-down system). Similarly, if the master IMU is not inertially fixed, a time-varying transformation, $M^{(o)}(t)$, is specified in terms of the body angles $\theta_\alpha^{(o)}(t)$, $\alpha = 1,2,3$, the rotations of the master IMU's relative orientations, as specified by $M_o^{(o)}$. Then, $T^{(i,\cdot)} = M_o^{(\cdot)} M(t)^{(\cdot)}$, and a vector in inertial coordinates, $\underline{v}_I$, can be expressed as a vector in moving (or body) coordinates $\underline{v}_B$ by the transformation $\underline{v}_B^{(\cdot)} = T^{(i,\cdot)} \underline{v}_I^{(\cdot)}$. Thus, we may write the thrust acceleration as $\underline{a}_B - \Delta\underline{a}_B = (T - \Delta T)(\underline{a}_I - \Delta\underline{a}_I)$ where the superscripts have been temporarily dropped for ease of notation, $\underline{a}_B$ is the true thrust acceleration in body coordinates, $\Delta\underline{a}_B$ is the error in measuring the thrust acceleration, expressed in body coordinates, and ΔT is the error in knowing the true transformation T.

Expanding the above equation and then neglecting the second-order terms gives

$$\Delta\underline{a}_B = \underline{a}_B - T\underline{a}_I + T\Delta\underline{a}_I + \Delta T\underline{a}_I$$

or

$$\Delta\underline{a}_I = T^{-1}\Delta\underline{a}_B - T^{-1}\Delta T\underline{a}_I \qquad (2.7)$$

Noting that the term $\Delta T T^{-1}$ can be written as an antisymmetric matrix of misalignment angles expressed in body coordinates,

$$\Delta T T^{-1} = [\Psi_B] = \begin{bmatrix} 0 & -\Psi_z & \Psi_y \\ \Psi_z & 0 & -\Psi_x \\ -\Psi_y & \Psi_x & 0 \end{bmatrix}, \quad \underline{\Psi}_B = [\Psi_x \ \Psi_y \ \Psi_z]^*,$$

we have from Equation (2.7)

$$\Delta \underline{a}_I = T^{-1} (\Delta \underline{a}_B - [\Psi_B] \ \underline{a}_B)$$

Now the transformation T is computed by integrating the equation

$$d(T - \Delta T)/dt = ([\omega_B] - [\varepsilon_B]) (T - \Delta T) \qquad (2.8)$$

where

$$[\omega_B] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix} \text{ and } \underline{\omega}_B = [\omega_x, \omega_y, \omega_z]^*$$

is the nominal angular velocity of the body, and where

$$[\varepsilon_B] = \begin{bmatrix} 0 & -\varepsilon_z & \varepsilon_y \\ \varepsilon_z & 0 & -\varepsilon_x \\ -\varepsilon_y & \varepsilon_x & 0 \end{bmatrix}, \quad \varepsilon_B = [\varepsilon_x, \varepsilon_y, \varepsilon_z]^* \text{ is a vector of}$$

gyro drift rates expressed in body coordinates. Expanding (2.8) and neglecting second-order terms gives

$$d(\Delta T)/dt = [\varepsilon_B]\, T - [\omega_B]\, \Delta T$$

but from the definition of $\Delta T T^{-1}$ we can write

$$(dT/dt)\,[\Psi_I] + T\,(d[\Psi_I]/dt) = [\varepsilon_B]\, T - [\omega_B]\, T\,[\Psi_I]$$

Next, we note that $(dT/dt)\,[\Psi_I] = -[\omega_B].\, T\,[\Psi_I]$ which implies

$$T(d[\Psi_I]/dt) = [\varepsilon_B]\, T \quad \text{or} \quad d[\Psi_I]/dt = T^{-1}\,[\varepsilon_B]\, T = [\varepsilon_I]$$

Using vector notation, this latter equation becomes

$$d\underline{\Psi}_I/dt = T^{-1}\, \underline{\varepsilon}_B = \underline{\varepsilon}_I$$

whose solution is

$$\underline{\Psi}_B(t) = T(t)\, \underline{\Psi}(t_o)_I + T\int_{t_o}^{t} T(\sigma)^{-1}\, \underline{\varepsilon}_B(\sigma)\, d\sigma \qquad (2.9)$$

This means Equation (2.7) may be written as

$$\Delta\underline{a}_I = T^{-1}(\Delta\underline{a}_B - T\,[\Psi_I]\, T^{-1}\, \underline{a}_B)$$

or, by defining an antisymmetric matrix $[\underline{a}_I]$ whose elements are the components of $\underline{a}_I$, we can write

$$\Delta\underline{a}_I = T^{-1}\, \Delta\underline{a}_B + [\underline{a}_I]\, \underline{\Psi}_I \qquad (2.10)$$

Substituting (2.9) into (2.10) gives

$$\Delta\underline{a}_I = T^{-1}\Delta\underline{a}_B + [a_I]\, \underline{\Psi}(t_o)_I + [a_I]\int_{t_o}^{t} T^{-1}(\sigma)\, \underline{\varepsilon}_B(\sigma)\, d\sigma \qquad (2.11)$$

From equations (2.22) and (2.23), infra, we have

$$\Delta \underline{a}_B = B_a(t)\,\underline{K}_a \quad \text{and} \quad \underline{\varepsilon}_B = B_g(t)\,\underline{K}_g$$

where $\underline{K}_a$ is a constant vector of accelerometer errors and $\underline{K}_g$ is a constant vector of gyro errors.

By defining a vector $\underline{K} = [\underline{\psi}(t_o)_I,\ \underline{K}_a,\ \underline{K}_g]^*$ and a matrix

$$B(t) = \left[[\underline{a}_I]\, T^{-1}(t)\, B_a(t) \,|\, [\underline{a}_I] \int_{t_o}^{t} T(\sigma)^{-1} B_g(\sigma)\, d\sigma \right]$$

we can write (2.11) as $\Delta \underline{a}_I = B(t)\underline{K}$. Thus, for the two IMU's

$$\Delta \underline{a}_I^{(\cdot)} = B^{(\cdot)}(t)\,\underline{K}^{(\cdot)} \tag{2.12}$$

where $\Delta \underline{a}_I^{(\cdot)}$ is a vector of errors in thrust acceleration due to the $IMU_{(\cdot)}$ instrument errors.

2.3 IMU Instrument Errors

The accelerometer measurement error for the i^{th} accelerometer may be written as

$$\Delta a_{m_i} = \nabla_i + [K_{a_i}, \alpha_{N_i}, \alpha_{C_i}]\,\underline{a}_T + \underline{a}_T^*\,[K_{g_i}^2]\,\underline{a}_T \tag{2.13}$$

and the gyro drift rate for the i^{th} gyro may be written as

$$\varepsilon_i = \varepsilon_{o_i} + [K_{\omega_i}, \gamma_{S_i}, \gamma_{o_i}]\ \underline{\omega} + [K_{I_i}, K_{S_i}, 0]\ \underline{a}_T$$

$$+ \underline{\omega}^* [K^2_{\omega_i}]\ \underline{\omega} + \underline{a}^*_T\ [K^2_{g_i}]\ \underline{a}_T \qquad (2.14)$$

In these equations, $\underline{a}_T$ is the thrust acceleration in the i^{th} accelerometer (gyro) coordinates; ∇_i is the bias error for the i^{th} accelerometer; K_{a_i} is the scale factor in the i^{th} accelerometer; α_{N_i} is the misalignment of the i^{th} accelerometer's input axis in the plane of the input and normal axes; α_{C_i} is the misalignment of the i^{th} accelerometer's input axis in the plane of the input and cross axes; $\underline{\omega}$ is the angular velocity in the i^{th} gyro coordinates (for ideal gimbaled systems this term will be the zero vector); ε_{o_i} is the i^{th} gyro's constant drift rate; K_{ω_i} is the scale factor of the i^{th} gyro; γ_{S_i} (γ_{o_i}) is the misalignment of the i^{th} gyro's input axis in the plane defined by the input and spin (output) axes; K_{I_i} (K_{S_i}) is the i^{th} gyro's error coefficient due to mass unbalance along the input (spin) axis; $\left[K^2_{g_i}\right]$ is a symmetric matrix where the elements $K^{(i)}_{jj}$ are the i^{th} accelerometers non-linearity for accelerations along the i^{th} axis and the elements $K^{(i)}_{jk}$ are the cross-axis sensitivity of the i^{th} accelerometer in the j - k plane; $\left[K^2_{\omega_i}\right]$ is a symmetric matrix where $K^{(i)}_{\omega_{jk}}$ are the error coefficients sensitive to the angular velocity components $\omega_j \omega_k$ for the i^{th} gyro; and $\left[K^2_{g_i}\right]$ is a symmetric matrix whose components $K^{(i)'}_{g_{jk}}$ are the error coefficients sensitive to the acceleration components $a_j a_k$ (anisolastic effect) for the i^{th} gyro.

It is certain that many of the elements in the above matrices will be so small that they may be neglected. If the gyros are used on a gimbaled platform, the terms proportional to ω and ω^2 would be neglected. By inserting the superscripts o and s in (2.13) and (2.14) the corresponding equations are obtained for the master and slave IMU's, respectively.

In the above error equations, the input accelerations and angular velocities are assumed to have been transformed into the coordinates of the particular instrument in question. Thus assuming a vector $\underline{v}_I^{(\cdot)}$ is expressed in some inertial frame, it may be expressed in accelerometer coordinates by

$$\underline{v}_{I_{acc}}^{(\cdot)} = A_I^{(\cdot)} T^{(i,\cdot)} \underline{v}_I^{(\cdot)} \equiv \mathcal{A}^{(\cdot)} \underline{v}_I^{(\cdot)}$$

or in gyro coordinates by replacing $A_I^{(\cdot)}$ by $G_I^{(\cdot)}$, so that

$$\underline{v}_{I_{gyro}}^{(\cdot)} = G_I^{(\cdot)} T^{(i,\cdot)} \underline{v}_I^{(\cdot)} \equiv \mathcal{L}^{(\cdot)} \underline{v}_I^{(\cdot)}$$

Applying these transformations to (2.13) and (2.14) gives

$$\Delta a_{m_i}^{(\cdot)} = \nabla_i^{(\cdot)} + [K_{a_i}^{(\cdot)}, \alpha_{N_i}^{(\cdot)}, \alpha_{C_i}^{(\cdot)}] \mathcal{A}^{(\cdot)} \underline{a}_T^{(I)} + \underline{a}_T^{(I)*} \mathcal{A}^{(\cdot)*} [K_{g_i}^{2(\cdot)}] \mathcal{A}^{(\cdot)} \underline{a}_T^{(I)} \tag{2.15}$$

and

$$\begin{aligned} \epsilon_i^{(\cdot)} = \epsilon_{o_i}^{(\cdot)} + [K_{\omega_i}^{(\cdot)}, \gamma_{S_i}^{(\cdot)}, \gamma_{o_i}] \mathcal{L}^{(\cdot)} \underline{\omega}^{(I)} + [K_{I_i}^{(\cdot)}, K_{S_i}^{(\cdot)}, 0] \mathcal{L}^{(\cdot)} \underline{a}_T^{(I)} \\ + \underline{\omega}^{(I)*} \mathcal{L}^{(\cdot)*} [K_{\omega_i}^{2(\cdot)}] \mathcal{L}^{(\cdot)} \underline{\omega}^{(I)} + \underline{a}_T^{(I)*} \mathcal{L}^{(\cdot)*} [K_{g_i}^{2'(\cdot)}] \mathcal{L}^{(\cdot)} \underline{a}_T^{(I)} \end{aligned} \tag{2.16}$$

when $\underline{a}_T^{(I)}$ is the thrust acceleration in some inertial reference frame and $\underline{\omega}^{(I)}$ is the angular velocity in the same inertial reference frame.

To express the acceleration measurement errors and gyro drift errors in platform coordinates, the components Δa_{m_i} and ε_i are multiplied by the transformation $\underline{A}_i^{-1}$ and $\underline{G}_i^{-1}$ respectively, thus

$$\Delta \underline{a}_m^{(\cdot)} = \left[\underline{A}_1^{(\cdot)-1} \vdots \underline{A}_2^{(\cdot)-1} \vdots \underline{A}_3^{(\cdot)-1}\right] \left[\Delta a_{m_1}^{(\cdot)}, \quad \Delta a_{m_2}^{(\cdot)}, \quad \Delta a_{m_3}^{(\cdot)}\right]^* \quad (2.17)$$

and

$$\underline{\varepsilon}^{(\cdot)} = \left[\underline{G}_1^{(\cdot)-1} \vdots \underline{G}_2^{(\cdot)-1} \vdots \underline{G}_3^{(\cdot)-1}\right] \left[\varepsilon_1^{(\cdot)}, \quad \varepsilon_2^{(\cdot)}, \quad \varepsilon_3^{(\cdot)}\right]^* \quad (2.18)$$

where the bar under A, G indicates that only the appropriate elements of $A_i^{(\cdot)}$ and $G_i^{(\cdot)}$ are included here.

If (2.15) is expanded, it could be written as the scalar product of a vector, $\underline{b}_{a_i}^{(\cdot)}$, which depends only on the nominal trajectory parameters and a vector of the i^{th} accelerometer error parameters, i.e.,

$$\Delta a_{m_i}^{(\cdot)} = \langle \underline{b}_{a_i}^{(\cdot)}, \underline{K}_{a_i}^{(\cdot)} \rangle \quad (2.19)$$

Similarly,

$$\varepsilon_i^{(\cdot)} = \langle \underline{b}_{g_i}^{(\cdot)} \quad \underline{K}_{g_i}^{(\cdot)} \rangle \quad (2.20)$$

Thus (2.17) and (2.18) may be written, respectively, as

$$\Delta \underline{a}_m^{(\cdot)} = [B_a^{(\cdot)}(t)]\, \underline{K}_a^{(\cdot)} \tag{2.21}$$

and

$$\underline{\varepsilon}^{(\cdot)} = [B_g^{(\cdot)}(t)]\, \underline{K}_g^{(\cdot)} \tag{2.22}$$

Letting

$$\underline{K}_a^{(\cdot)} = \left[\underline{K}_{a_1}^{(\cdot)}, \underline{K}_{a_2}^{(\cdot)}, \underline{K}_{a_3}^{(\cdot)}\right], \quad \underline{K}_g^{(\cdot)} = \left[\underline{K}_{g_1}^{(\cdot)}, K_{g_2}^{(\cdot)}, K_{g_3}^{(\cdot)}\right]^* \quad \text{where}$$

$$K_{a_i}^{(\cdot)} = \left[\nabla_i^{(\cdot)}, K_{a_i}^{(\cdot)}, \alpha_{n_i}^{(\cdot)}, \alpha_{s_i}^{(\cdot)}, K_{x_x}^{(\cdot)}, \ldots\right]^* \tag{2.23}$$

$$K_{g_i}^{(\cdot)} = \left[\varepsilon_{o_i}^{(\cdot)}, K_{\omega_i}^{(\cdot)}, \gamma_{s_i}^{(\cdot)}, \gamma_{c_i}^{(\cdot)}, K_{\omega_{xx}}^{(\cdot)}, \ldots, K_{g_{xx}}^{(\cdot)}, \ldots\right]^* \tag{2.24}$$

and the matrices $[B_a^{(\cdot)}(t)]$, $[B_g^{(\cdot)}(t)]$ are defined by equations (2.17), (2.18), (2.19), and (2.20).

In any realistic problem, a good many of the components in (2.13) and (2.14) would be zero, especially for the master IMU and for quadratic error terms. Equations (2.21) and (2.22) are a useful form of the instrument error terms for use in the minimum-variance estimator since the vectors $\underline{K}_{a_i}^{(\cdot)}$, $\underline{K}_{g_i}^{(\cdot)}$ are constant vectors depending only on the error parameters of the accelerometers and gyros.

2.4 Structure of the Minimum-Variance Estimator

It is necessary to assume that a statistical description of random inputs due to vehicle induced, random acceleration is available, and that these inputs can be represented by white noise passing through a linear dynamical system. Such a linear filtering operation is performed by what is usually called a shaping filter (Appendix "A" presents the method by which the random vibrations, $\underline{\eta}(t)$, can be represented by a linear dynamical system along with illustrative examples). Thus, it is assumed that the vehicle induced, random accelerations may be represented by

$$d\underline{\eta}'/dt = A^p(t)\underline{\eta}' + B^p(t)\underline{w}(t) \qquad (2.25)$$

where $A^p(t)$ and $B^p(t)$ are, in general, time varying matrices which are related to the covariance matrix of $\underline{\eta}$, and $\underline{w}(t)$ is a white noise vector process. The prime on $\underline{\eta}$ is to indicate that $\underline{\eta}'$ may be composed of $\underline{\eta}$ and other additional suitably defined vectors arranged one upon the other. We shall write $\underline{\eta} = [I']\underline{\eta}'$ to indicate that only the components $\underline{\eta}$ are used from $\underline{\eta}'$. We now combine equations (2.6), (2.12), and (2.25) and get

$$\frac{d}{dt}\begin{bmatrix} \underline{p}_{meas}^{(o,s)} \\ \underline{v}_{meas}^{(o,s)} \\ \underline{\eta}^{(o,s)'} \\ \underline{k}^{(s)} \\ \underline{k}^{(o)} \\ \varepsilon \underline{b}_o^{(o,s)} \end{bmatrix} = \left[\begin{array}{c|c|c|c|c|c} 0 & I & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & [\Omega][\Gamma'] & -B^{(s)}(t) & B^{(o)}(t) & [\Omega'] \\ \hline 0 & 0 & A^{p}(t) & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \begin{bmatrix} \underline{p}_{meas}^{(o,s)} \\ \underline{v}_{meas}^{(o,s)} \\ \underline{\eta}^{(o,s)'} \\ \underline{k}^{(s)} \\ \underline{k}^{(o)} \\ \varepsilon \underline{b}_o^{(o,s)} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ B^{p}(t) \\ 0 \\ 0 \\ 0 \end{bmatrix} \underline{w} + \begin{bmatrix} 0 \\ [\Omega'] \underline{b}_o^{(o,s)} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

(2.26)

The observations, or measurements, are assumed to be $\underline{p}^{(o,s)}_{meas}$ and $\underline{v}^{(o,s)}_{meas}$. (Actually one might also include the vehicle's angular velocity as an observation, but one can show that the angular velocity equation can be uncoupled from the rest; thus, it can be studied independantly.) These observations are assumed to have errors, and it is assumed that the error in the observations, because of observation error parameters, may be written as

$$\Delta\underline{z} = B_m(t)\,\underline{k}^m \tag{2.27}$$

where $\Delta\underline{z}$ is a vector denoting the error in the observation vector, $\underline{z}$, due to the observation error parameters, $\underline{k}^m$, which are assumed to be constants over a particular observation period, and $B_m(t)$ is a matrix relating the observation error parameters to the nominal (or known) variables in the observation and dynamical equations. The vector $\underline{k}^m$ would include such things as biases and scale factor errors in the various IMU integrators.

Next the observations are assumed to be contaminated by a random noise $\underline{v}$. To be general, it is assumed that this noise could be correlated, so it is expressed as

$$\underline{v} = \underline{\nu}_c + \underline{v}_w \tag{2.28}$$

where $\underline{\nu}_c$ is a vector of correlated noise components and $\underline{v}_w$ is a

vector of white noise components. As above (2.25), it is assumed the correlated noise may be represented in the Markovian form

$$d\underline{\nu}'_c/dt = A^m(t)\underline{\nu}'_c + B^m(t)\underline{\nu}'_c \tag{2.29}$$

where the prime carries the same connotation as above. The observations are thus

$$\underline{z} = \begin{bmatrix} \underline{p}_{meas} \\ \underline{v}_{meas} \end{bmatrix} + B_m(t)\underline{k}^m + [I']\underline{\nu}'_c + \underline{v}_w \tag{2.30}$$

Defining the state vector as

$$\underline{X} = \left(\underline{p}, \underline{v}, \underline{\pi}', \underline{\nu}'_c, \underline{k}^{(s)}, \underline{k}^{(o)}, \underline{k}^m, \varepsilon\underline{b}^{(o,s)}\right)^*$$

Equations (2.26) and (2.30) may be written in canonical form as

$$\begin{aligned} \dot{\underline{X}} &= AX + B\underline{W} + \underline{f} \\ \underline{Z} &= H\underline{X} + \underline{V} \end{aligned} \tag{2.31}$$

where

$$A = \left|\begin{array}{c|c|c|c|c|c|c|c} 0 & I & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & [\Omega I'] & 0 & -B^{(s)} & B^{(o)} & 0 & [\Omega] \\ \hline 0 & 0 & A^p & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & A^m & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right|, \quad B = \left|\begin{array}{c|c} 0 & 0 \\ \hline 0 & 0 \\ \hline B^p & 0 \\ \hline 0 & B^m \\ \hline 0 & 0 \end{array}\right|$$

$$H = [I \mid 0 \mid I' \mid 0 \mid 0 \mid B_m \mid 0], \qquad \underline{W} \equiv (\underline{w}, \underline{\nu}'_c)^*$$

$$\underline{f} = \left(\underline{0}, [\Omega']\underline{b}_o^{(o,s)}, \underline{0}, \underline{0}, \underline{0}, \underline{0}, \underline{0}\right)^*, \quad \underline{V} \equiv \underline{v}_w$$

Equations (2.31) have the discrete form $\underline{X}(k+1) = \phi(k+1,k)\underline{X}(k) + \underline{W}(k)$ and $\underline{Z}(k+1) = H(k+1)\underline{X}(k+1) + \underline{V}(k+1)$. We next show how the discrete equation set is formulated in terms of a specific problem.

2.5 Specialized Calibration and Alignment Problem

In this concluding section a discrete form of (2.31) is developed in a manner to give concrete meaning to the ideas presented above, and to indicate the approach used in the next chapters where further examples and numerical results are presented. In the problem considered here, IMU_0 is well calibrated and aligned prior to vehicle operation so that IMU_s could be considered adequately calibrated and aligned when compared with IMU_0 as a standard. Thus, the vector of IMU_0 error parameters, $k^{(o)}$, may be taken to be the zero vector.

Next it is assumed nonlinear error terms and fixed misalignment angles on the accelerometers and gyros are negligible and that these instruments are nominally located so that the instruments' sensitive axes form orthogonal triads. Under these assumptions the matrix $B^{(s)}(t)$ becomes

$$B^{(s)}(t) = \left[[a] \mid M^{(s)}(t)\ M^{(s)}\ \mathcal{A} \mid [a] \int_{t_0}^{t} d\sigma\ M^{(s)}(\ )\ M^{(s)}\ \mathcal{B} \right]$$

a 3x21 matrix where

$$\mathcal{A} = \begin{bmatrix} 1 & 0 & 0 & a_x & 0 & 0 \\ 0 & 1 & 0 & 0 & a_y & 0 \\ 0 & 0 & 1 & 0 & 0 & a_z \end{bmatrix}$$

$$\mathcal{B} = \begin{bmatrix} 1 & 0 & 0 & \omega_x & 0 & 0 & a_x & a_y & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \omega_y & 0 & 0 & 0 & a_y & a_z & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \omega_z & 0 & 0 & 0 & 0 & a_z & a_x \end{bmatrix}$$

and $\underline{k}^{(s)}$ is given by the 21-vector

$$\underline{k}^{(s)*} = \left[\underline{x}^{(s)}(t_o),\ \underline{\Delta}^{(s)},\ \underline{k}_a^{(s)},\ \underline{\varepsilon}_o^{(s)},\ \underline{k}_\omega^{(s)},\ \underline{\alpha}_\omega^{(s)} \right]$$

In the matrix $B^{(s)}(t)$, a_x, a_y, a_z, ω_x, ω_y, ω_z are nominal acceleration and angular velocity components in the IMU_s coordinates. Assuming the random angular acceleration of the vehicle is negligible, $\underline{\omega}$ small or nearly constant, means the "angular velocity matrix" is defined by $[\Omega] = \left| [I] \mid 2\,[\omega] \right|$, and by defining a vehicle vibration vector as $\underline{\eta}(t)^* = [\underline{\eta}\ \dot{\underline{\eta}}]$, then (2.6) becomes

$$\frac{d}{dt} \begin{bmatrix} \underline{p}_m \\ \underline{v}_m \end{bmatrix} = \left[\begin{array}{c|c} 0 & I \\ \hline 0 & 0 \end{array}\right] \begin{bmatrix} \underline{p}_m \\ \underline{v}_m \end{bmatrix} + \left[\begin{array}{c} 0 \\ \hline -\Delta\underline{a}_m^{(s)} + \ddot{\underline{\eta}} + [\omega]\,\dot{\underline{\eta}} \end{array}\right]$$

Next it is assumed $\underline{\eta}$ is suitably approximated by

$$d\underline{n}/dt = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \underline{n} + B'(t)\,\underline{w}$$

where $\underline{n}^* = [\underline{n}_2, \underline{n}_3]$ where $\underline{n}_2$ and $\underline{n}_3$ are 3-vectors and the A_{ij} are 3x3 time varying matrices, $\underline{n}_2 = \ddot{\underline{\eta}}$, $\underline{n}_1 = \dot{\underline{\eta}} = \int \underline{n}_2 \, dt$. $B'(t)$ is a 6x6 time varying matrix and $\underline{w}$ is a 6-white noise vector. Thus we can write $\dot{\underline{\eta}}' = A^p(t)\,\underline{\eta}' + B^p(t)\,\underline{w}$ where $\underline{\eta}'^* \equiv [\underline{n}_1, \underline{n}_2, \underline{n}_3] = [\dot{\underline{\eta}}, \ddot{\underline{\eta}}, \underline{n}_3]$ and

$$A^p(t) \equiv \begin{bmatrix} 0 & [I'] & 0 \\ 0 & A_{11} & A_{12} \\ 0 & A_{21} & A_{22} \end{bmatrix}, \quad B^p(t) \equiv \begin{bmatrix} 0 \\ B'(t) \\ 0 \end{bmatrix}, \quad [I'] = [[I] \,|\, 0]$$

where [I] is a 6x6 identity matrix, $\underline{w}$ is uncorrelated, and $\underline{k}^{(m)} = \underline{0}$. Assuming the angular velocity equation is treated separately, the observation equation becomes $\underline{Z} = \begin{bmatrix} \underline{p}_m \\ \underline{v}_m \end{bmatrix} + \underline{v}(t)$ where $\underline{v}(t)$ is a 6-vector of white measurement noise. If we define a 36-state vector as $\underline{X}^* = \left[\underline{p}_m, \underline{v}_m, \ddot{\underline{\eta}}, \dot{\underline{\eta}}, \underline{n}_3, \underline{k}^{(s)}\right]$ and the matrices

$$A(t) \equiv \begin{bmatrix} 0 & I & 0 & 0 \\ 0 & 0 & M^{(m)} [\,I \,|\, [\omega]\,] \; [I'] & -B^{(s)} \\ 0 & 0 & A^p & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B(t) \equiv \begin{bmatrix} 0 \\ B^p \\ 0 \end{bmatrix}$$

and $H = \left[[I] \,|\, 0\right]$, then the continuous problem is specified by $\dot{\underline{X}} = A(t)\,\underline{X} + B(t)\,\underline{w}$ and $\underline{Z} = H\underline{X} + \underline{v}(t)$. The equations for the discrete formulation are

$$\underline{X}_{k+1} = \phi(k+1,k)\,\underline{X}_k + \underline{w}_k \quad \text{and} \quad \underline{Z}_{k+1} = H(k+1)\underline{X}(k+1) + \underline{v}_{k+1} \tag{2.32}$$

where $\Phi(k+1,k)$ is given by

$$\Phi(k+1,k) \equiv \left[\begin{array}{c|c|c} \Phi_1(k+1,k) & \Phi_2(k+1,k) & \Phi_3(k+1,k) \\ \hline 0 & \Phi_4(k+1,k) & 0 \\ \hline 0 & 0 & I_1 \end{array}\right] \quad (2.33)$$

The submatrices of (2.33) are given by

$$\Phi_1(k+1,k) = \left[\begin{array}{c|c} I & \Delta t \cdot I \\ \hline 0 & I \end{array}\right]$$

$$\Phi_1'(k+1,k) = \left[\begin{array}{c} \Delta t\, I \\ \hline I \end{array}\right]$$

$$\Phi_2(k+1,k) = \int_k^{k+1} \Phi_1'(\sigma,k)\, M^{(m)}(\sigma)\,[[I]\,|\,[\omega]]\, d\sigma\,[I']$$

$$\Phi_3(k+1,k) = -\int_k^{k+1} \Phi_1'(\sigma,k) B^{(s)}(\sigma)\, d\sigma$$

I_1 is a 21x21 identity matrix and $\Phi_4(k+1,k)$ is the solution of the homogeneous equation $d\,\Phi_4(t,\tau)/dt = A^p(t)\,\Phi_4(t,\tau)$ where, as indicated below, $\Phi_4(t,t) = I$. Qualitatively, Φ_2 is a matrix which gives the contributions to $\underline{p}_m$ and $\underline{v}_m$ due to vehicle vibrations $\underline{\eta}$, and Φ_4 is the 9x9 transition matrix for the random vehicle vibrations. The white noise sequence $\underline{w}_k$ is given by $\underline{w}_k = \int_k^{k+1} \left[\Phi_4(\sigma, k) B^p(\sigma)\, \underline{w}(\sigma)\right] d\sigma$. The white noise observation sequence $\underline{v}_{k+1}$, is similarly obtained from the white noise observation process, $\underline{v}(t)$. I_1 = identity · indicates that the error

parameters $\underline{k}^{(s)}$ remain constant. The transition matrices have the following properties:

1) $\phi(t_3, t_2)\phi(t_2, t_1) = \phi(t_3, t_1)$ for all t_1, t_2, t_3,

2) $\phi(t_2, t_1)^{-1} = \phi(t_1, t_2)$,

3) $\phi(t, r) = X(t)X(\tau)^{-1}$ where $X(t)$ is the solution to the equation $dX/dt = A(t)X(t)$, $X(o) = I$.

From the above properties it is seen that if $t_o = 0$, $\phi(t_1, 0) = X(o)$, and it is not difficult to see that in general $\phi(t_{k+1}, t_k) = X^{(k+1)}(t_1) \equiv \phi^{(k+1)}(t_1, 0)$ where $X^{(k+1)}(t_k)$ is the solution to the matrix equation $\dot{X}(t) = A(t+t_k)X(t)$, $X(o) = I$. This means that it is only necessary to solve for the matrix $X(t_1)$ for each time interval, (t_k, t_{k+1}), with an updated $A(t)$ matrix using the same initial conditions $X(o) = I$. Thus the computer routine for generating the transition matrix, $\phi(t_{k+1}, t_k)$ is not changed for each k, only the elements of $A(t)$ need to be updated.

CHAPTER III

FORMULATION OF SPECIFIC ERROR PARAMETER IDENTIFICATION PROBLEMS

In Chapter I, a very general formulation of the IMU error parameters was considered. This formulation is general in the sense that most specific problems can be placed into this framework. When a specific problem is considered, the equations of Chapter II simplify considerably and the dimensions of the various matrices are decreased accordingly. For example, it could be assumed that all the error parameters are known and that only the initial misalignment angles, $\underline{\Psi}(t_0)$, need to be estimated. Another important simple example might be to estimate the random motion only. This might occur when a vehicle is launched from the wing of an airplane or mother-ship.

In this chapter we discuss three important IMU error parameter configurations which were considered in this study. They are as follows:

1. Initial misalignment angles, $\underline{\Psi}$, only.
2. Initial misalignment angles plus gyro mass-unbalance terms, $\underline{k}_S$ and $\underline{k}_I$.
3. Initial misalignment angles plus gyro constant drift rates, $\underline{\varepsilon}$.

Both time correlated and uncorrelated random accelerations are considered. The correlated model used is discussed in considerable detail in Appendix "A". The correlated random acceleration is assumed to act along one axis of the vehicle (axis of major disturbance) and depends upon specific parameters which cannot be readily varied for parametric or analytic studies. For this reason extensive experimental results for this model are not included - although they are available. The main purpose of this study is to determine effects of state and observation noise levels, and the trajectory parameters on the identification of the errors, and the uncorrelated noise models are most convenient for these purposes.

For these specific models, the master IMU is assumed to be perfect, and the slaved IMU is assumed to be a gimballed system. The only degree of freedom in the nominal trajectory specification is assumed to be in the vehicle's pitch angle. That is, the vehicle is assumed to have a fixed thrust program. The experimental results of this chapter indicate that the quality of the identification depends very much on the manner in which the particular error parameters enter the system and on the power levels in the random disturbances. Thus a meaningful optimization problem can be formulated with respect to the nominal trajectory variables, such as for example, the trajectory pitch angle.

3.1 Time Correlated Random Vehicle Accelerations

For the present purposes it is assumed that significant random motion occurs along the vehicle's yaw axis only. Such a random motion might occur if the thrust engine were to move randomly about the pitch axis while controlling the vehicle's trajectory to lie in a vertical plane. Furthermore, it is assumed that this random motion is stationary and has a power spectrum of the form discussed in Appendix "A", Section A.1. In this section the assumption that the given power spectrum may be suitable approximated by the fourth-order rational power spectrum

$$S(\omega) = \frac{\Phi_o b}{(\omega - a)^2 + b^2} + \frac{\Phi_o b}{(\omega + a)^2 + b^2} \qquad (3.1a)$$

is made. The corresponding autocorrelation function is

$$R_x(\tau) = \Phi_o \exp(-b|\tau|) \cos a\tau \qquad (3.1b)$$

By Equation (A.5) of Appendix A, a model of the process defined by Equation (3.1a) is given by

$$\frac{d}{dt}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -b & a \\ -a & -b \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \sqrt{2b\Phi_o}\begin{pmatrix} 1 \\ \frac{c-b}{a} \end{pmatrix} w, \quad c \equiv \sqrt{a^2 + b^2} \qquad (3.1c)$$

The "noise", $w(t)$, is such that $E[w(t)] = E[w(t)\,w(\tau)] = \delta(t - \tau)$. For engineering purposes, Equation (3.1c) is usually discretized by

approximating the white noise process $w(t)$ by the white noise sequence $w_k/\sqrt{\Delta}$ where k denotes the number of the sample and Δ denotes the sampling interval (see Section A.4). The discrete form of (3.1c) is given by (A.20c) as

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = e^{-b\Delta} \begin{pmatrix} \cos a\Delta & \sin a\Delta \\ -\sin a\Delta & \cos a\Delta \end{pmatrix} \begin{pmatrix} x_i - 1 \\ y_i - 1 \end{pmatrix} + \tag{3.1d}$$

$$+ \frac{1}{c}\sqrt{\frac{2\Phi_o}{b\Delta}} \begin{pmatrix} e^{-b\Delta}\left(\frac{c-b}{a}\sin a\Delta - \cos a\Delta\right) + 1 \\ e^{-b\Delta}\left(\frac{c-b}{a}\cos a\Delta + \sin a\Delta\right) + \left(\frac{b-c}{a}\right) \end{pmatrix} w_{i-1}$$

Simplifying the notation, Equation (3.1d) may be written as

$$\underline{x}_i = \underline{A}_{II}\, x_{i-1} + \underline{b}_{II}\, w_{i-1} \tag{3.1}$$

where the definition of $\underline{x}_i$, $\underline{A}_{II}$, and $\underline{b}_{II}$ is obvious from Equation (3.1c). The "white noise" sequence w_{i-1} has the property

$$E\left[w_i\right] = 0 \text{ and } \operatorname{cov}(w_i, w_j) = E\left[(w_i\, w_j)\right] = \delta_{ij} \tag{3.1e}$$

Another discrete model of the process defined by Equation (3.1a) is given in Section A.4, Equation (A.21c), as

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} b_1 & -b_2 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix} + \begin{pmatrix} a_1 \\ \frac{-a_2}{b_2} \end{pmatrix} w_{i-1} \tag{3.2a}$$

To simplify the notation, Equation (3.2a) is written

$$\underline{x}_i = A_{III}\, \underline{x}_{i-1} + \underline{b}_{III}\, w_{i-1} \tag{3.2b}$$

The definitions of a_1, a_2, b_1, b_2 are given in Section A.4.

Next it is assumed that the random motion which is introduced through the "angular velocity matrix" $[\Omega']$ is negligible. This is a reasonable assumption if the nominal angular velocity $\underline{\omega}$ is small and nearly constant. In the present case, $\underline{\omega}^{(m)} = (0, \dot{\theta}_2^{(o,m)}, 0)^*$ where $\theta_2^{(o,m)}$ is the vehicle pitch angle and $\dot{\theta}_2^{(o,m)}$ is zero except at the transitions where it is a constant. This means that the "angular velocity matrix" $[\Omega]$ is defined by

$$[\Omega] = \left[I \middle| 2\,[\omega]\right] = \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & -2\omega_3 & 2\omega_2 \\ 0 & 1 & 0 & 2\omega_3 & 0 & -2\omega_1 \\ 0 & 0 & 1 & -2\omega_2 & 2\omega_1 & 0 \end{array}\right] =$$

$$= \left[\begin{array}{ccc:ccc} 1 & 0 & 0 & 0 & 0 & 2\omega_2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2\omega_2 & 0 & 0 \end{array}\right] \qquad (3.3a)$$

and the vehicle's random acceleration vector is

$$\underline{\Pi}(t) = \left[0 \;\; 0 \;\; \ddot{n} \;\; 0 \;\; 0 \;\; \mathring{n}\right]^* \qquad (3.3b)$$

According to Equation (3.1c), $\mathring{n}$ and $\ddot{n}$ may be expressed in the form

$$\frac{d}{dt}\begin{pmatrix} n \\ \dot{n} \\ \mathring{n} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} \ddot{n} \\ \dot{n} \\ \mathring{n} \end{pmatrix} + \begin{pmatrix} b_{11} \\ b_{21} \\ 0 \end{pmatrix} w \qquad (3.3c)$$

where $\mathring{n}'$ is an additional variable which is required to take care of

the correlation in $\ddot{\underline{n}}$. The elements in the matrix equation above are obtained from Equations (3.1c) or (3.2a). Equation (3.3c) may be written in discrete form as

$$\begin{pmatrix} \ddot{\underline{n}} \\ \ddot{\underline{n}}' \\ \dot{\underline{n}} \end{pmatrix}_{k+1} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & 0 \\ \alpha_{21} & \alpha_{22} & 0 \\ \Delta & 0 & 1 \end{pmatrix} \begin{pmatrix} \ddot{\underline{n}} \\ \ddot{\underline{n}}' \\ \dot{\underline{n}} \end{pmatrix}_{k} + \begin{pmatrix} \beta_{11} \\ \beta_{21} \\ 0 \end{pmatrix} w_k \qquad (3.3d)$$

where the subscripts k + 1, k refer to the sampling interval, and the elements in the matrices are obtained from Equation (3.1d) or (3.2b).

The random motion of the vehicle, in vehicle coordinates, may thus be written as

$$\underline{n}^{(m)} = [\Omega]\,[\dot{\Gamma}]\,\underline{\eta}' \equiv \left[I \,\middle|\, 2[\omega]\right]\,[I \;\; 0]\,[0 \;\; 0 \;\; \dot{n} \;\; 0 \;\; 0 \;\; \ddot{n} \;\; \ddot{\eta}']^*$$

If we substitute into the above matrices, this becomes

$$\underline{n}^{(m)} = \begin{pmatrix} 2\omega_2 & \dot{n} \\ -2\omega_1 & \dot{n} \\ & \ddot{\eta} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2\omega_2 \\ 0 & 0 & -2\omega_1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ddot{n} \\ \ddot{\eta}' \\ \dot{n} \end{pmatrix} \qquad (3.3e)$$

Equation (3.3e) may be written as

$$\underline{n}^{(m)} \equiv [\Omega''] \; \underline{\eta}' \qquad (3.4a)$$

where

$$[\Omega'] \equiv \begin{bmatrix} 0 & 0 & 2\dot{\theta}_2^{(o,m)} \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \underline{n}' = (\ddot{n}, \ddot{n}', \dot{n})^* \tag{3.4b}$$

In [3] it is shown that under reasonable assumptions the effects of $\underline{\theta}_2^{(o,m)}$ may be neglected; that is, random coriolis accelerations can be neglected.

3.2 Uncorrelated Random Accelerations

For various reasons, it may be possible to assume that the discrete form of the random acceleration due to vehicle vibrations may be uncorrelated in time. Such would be the case, for example, if the estimation interval Δ was required to be greater than, say 0.1 sec. Such a requirement would depend on how fast a digital computer could process the data, i.e., the velocity differences between the two IMU's. In the present study, it has been found that, in order to estimate initial misalignments and mass-unbalance drift rates, Δ would need to be greater than 0.1 sec. In the event that $\Delta > 0.1$ Figure 3.1 illustrates how the correlated noise model assumed in the previous section might generate essentially uncorrelated velocity sequences. As a further example, a sensible candidate

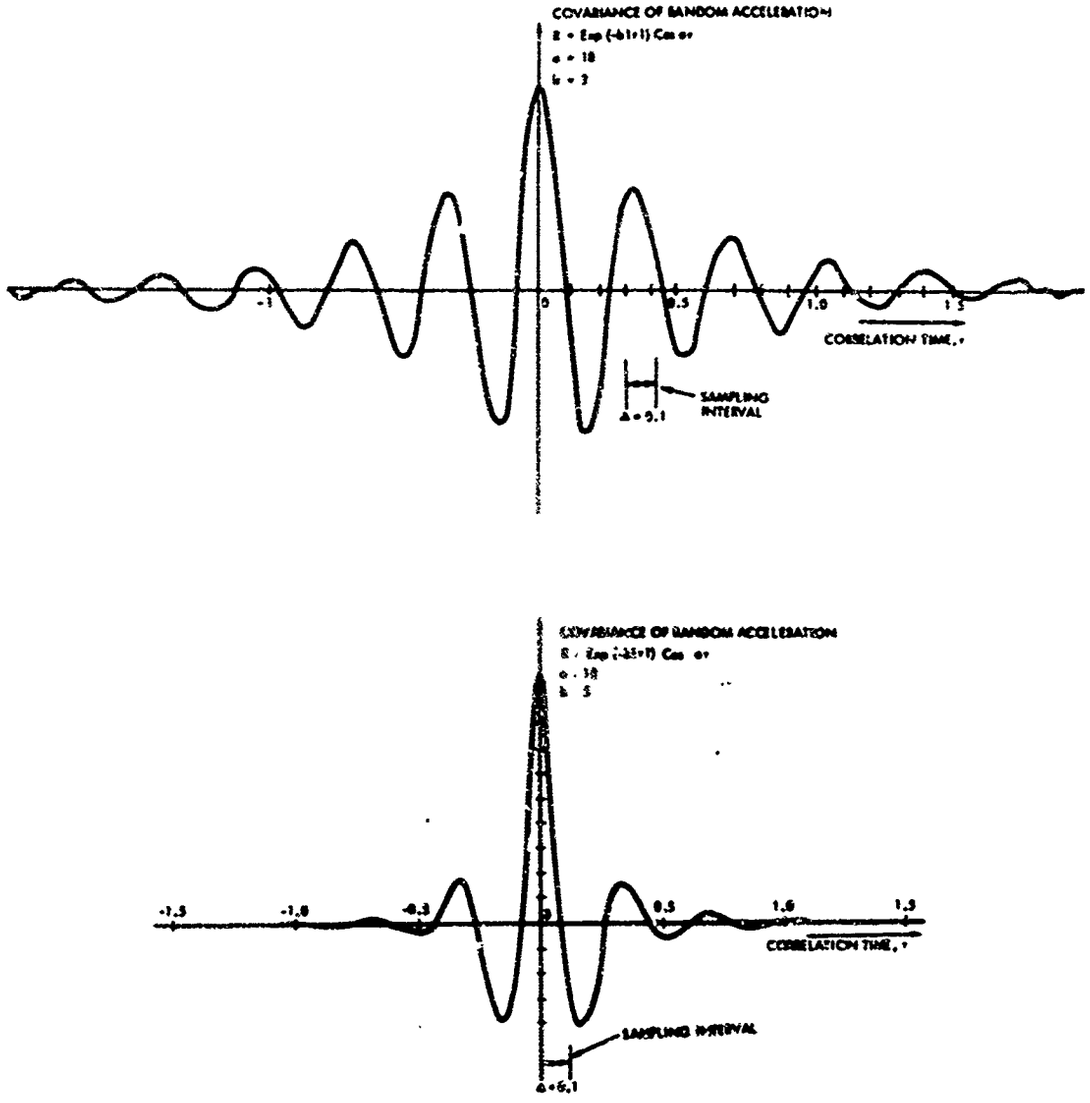


Figure 3.1 **Illustrations of the Band-Pass Correlated Noise Model**

for a suboptimal estimator would be one in which the random accelerations are assumed uncorrelated in time. It is thus assumed that the random acceleration added to the IMU velocity differences is of the form

$$\underline{r}_k \equiv \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}_k \quad ; \quad \text{where } E\left[\underline{r}\,\underline{r}^*\right]_k = \begin{bmatrix} r_{11} & 0 & 0 \\ 0 & r_{22} & 0 \\ 0 & 0 & r_{33} \end{bmatrix}_k \equiv R \qquad (3.5a)$$

Further, it is assumed that $\sigma_R^2 \equiv r_{11} = r_{22} = r_{33}$. This assumption is made for computational convenience and is not conceptual in nature. The random acceleration in IMU_s coordinates is given by $\left(\text{with } T \equiv T^{(o,s)}\ T^{(m,o)}\right)$,

$$\underline{r}_k^{(s)} = \int_{k\Delta}^{(k+1)\Delta} T^{(o,s)}\, T^{(m,o)}(t)\, dt\ \frac{\underline{r}_k^{(m)}}{\sqrt{\Delta}} \qquad (3.5b)$$

The covariance of this random acceleration is given by,

$$\text{Cov}(\underline{r}_k, \underline{r}_j) = E\left[\left(\int_{t_k}^{t_{k+1}} T(t)\,\underline{r}(t)\,dt\right).\left(\int_{t_j}^{t_{j+1}} T(\tau)\,\underline{r}(\tau)\,d\tau\right)\right] \qquad (3.5c)$$

$$= \sigma_R^2\, I$$

due to the fact that $T(t)$ is an orthogonal transformation.

3.3 Initial Misalignment Angles, Ψ_i, $i = 1, 2, 3$

If it is assumed that the accelerometer error parameters $\underline{k}_a^{(s)}$ and the gyro error parameters $\underline{k}_g^{(s)}$ are either zero or known (Chapter II), then the difference in acceleration between the slave and master IMU (expressed the IMU_s coordinates) becomes

$$\Delta\underline{a}^{(s,o)} = k_{mis} \left[\underline{a}^{(s)}\right] \underline{\Psi} + T^{(m,s)} [\Omega''] \underline{\eta}' \qquad (3.6a)$$

In this equation the terms are defined in Chapter II.

If the random coriolis acceleration is neglected (as described above) then

$$[\Omega''] \underline{\eta}' = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \qquad (3.6b)$$

In order to write Equation (3.6b) in discrete form, let

$$\underline{v}_k^{(s,o)} \equiv \underline{v}_{k-1}^{(s,o)} + \int_{k-1}^{k} \Delta\underline{a}^{(s,o)}(\tau)\, d\tau \qquad (3.6c)$$

where k denotes the time $t = t_k$. Equation (3.6c) may then be written as

$$\underline{v}_k^{(o,s)} = \underline{v}_{k-1}^{(o,s)} + k_{mis} \int_{k-1}^{k} \left[\underline{a}^{(s)}(\tau)\right] d\tau \underline{\Psi} + \int_{k-1}^{k} T^{(m,s)}(\tau) d\tau [\Omega''] \underline{\eta}_{k-1} \qquad (3.6d)$$

Combining Equations (3.6d) and (3.2b) into one equation gives

$$\begin{bmatrix} \underline{v}_k^{(o,s)} \\ \hline \underline{\eta}_k \\ \hline \underline{\Psi}_k \end{bmatrix} = \left[\begin{array}{c|c|c} I & \int_{k-1}^{k} T^{(m,s)}(\tau)\, d\tau\, [\Omega''] & k_{mis} \int_{k-1}^{k} \left|a^{(s)}(\tau)\right| d\tau \\ \hline 0 & A_{III} & 0 \\ \hline 0 & 0 & I \end{array}\right] \begin{bmatrix} \underline{v}_{k-1}^{(o,s)} \\ \hline \underline{\eta}_{k-1} \\ \hline \underline{\Psi}_{k-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \hline \underline{b}_{III} \\ \hline 0 \end{bmatrix} w_{k-1} \tag{3.6e}$$

which is the state equation for the misalignment problem.

It is assumed that the velocity increments between the two IMU's can be observed. However, these observations are contaminated by IMU instrument quantization errors and digital computer round-off errors. Further, it is assumed that these errors form an uncorrelated random sequence. Thus the observations made are

$$\begin{bmatrix} y_1^{(o,s)} \\ y_2^{(o,s)} \\ y_3^{(o,s)} \end{bmatrix} = \begin{bmatrix} v_1^{(o,s)} \\ v_2^{(o,s)} \\ v_3^{(o,s)} \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

where the w_i, $i = 1, 2, 3$, are uncorrelated sequences of gaussian random variables with

$$\text{cov}\,(\underline{w}_i, \underline{w}_j) = \underline{W}_i\ \delta_{ij} \tag{3.7b}$$

Next a state vector $\underline{x}_k$ is defined as

$$\underline{x}_k \equiv \left[\underline{v}_k^{(o,s)}, \underline{\eta}_k, \underline{\Psi}_k\right]^* \tag{3.8a}$$

an observation vector

$$y_k \equiv \left[y^{(o,s)},\ y^{(o,s)},\ y^{(o,s)} \right]_k \tag{3.8b}$$

a state transition matrix

$$\phi_{k,k-1} \equiv \left[\begin{array}{c|c|c} I & \int_{k-1}^{k} T^{(m,s)}(\tau)\, d\tau\, [\Omega''] & k_{mis} \int_{k-1}^{k} \left| a^{(s)}(\tau) \right| d\tau \\ \hline 0 & A_{III} & 0 \\ \hline 0 & 0 & I \end{array}\right] \tag{3.8c}$$

a measurement matrix

$$M = [I \mid 0] \ ; \ I \text{ is } 3\text{x}3 \text{ and } 0 \text{ is } 3\text{x}5 \tag{3.8d}$$

and a process noise vector

$$\underline{v}_k = \left[\begin{array}{c} 0 \\ \hline b_{III} \\ \hline 0 \end{array}\right] w_k \tag{3.8e}$$

With these definitions, Equations (3.6c) and (3.7a) become

$$\underline{x}_k = \phi_{k,k-1}\, \underline{x}_{k-1} + \underline{v}_{k-1} \tag{3.9}$$

$$\underline{y}_k = M\, \underline{x}_k + \underline{r}_k \tag{3.10}$$

with the covariance matrices of $\underline{w}_k$ and $\underline{r}_k$ given by W_k and R_k respectively.

In the discussion which follows, it is convenient to define the quantities

$$\Delta \underline{v}_{i.m.}(k+1,k) = k_{mis} \int_{k\Delta}^{(k+1)\Delta} \left(a^{(s)}(\tau) \right) d\tau \, \underline{\Psi} \qquad (3.11a)$$

and

$$\left(v_s^{(s)} \right)_{i.m.} = \left(v_s^{(s)}(k+1,k) \right)_{i.m.} = \int_{k\Delta}^{(k+1)\Delta} \left(a_s^{(s)}(\tau) \right) d\tau \qquad (3.11b)$$

3.4 Initial Misalignment Angles and Mass-Unbalance Drift Rates

In this section we obtain the equations for estimating initial gyro misalignment angles, Ψ_i, and the gyro drift terms which are due to gyro mass unbalances. The drift rate for the i^{th} gyro is

$$\varepsilon^{(g_i)} = k_{m.u.} \begin{bmatrix} k_{I_i} & k_{S_i} & 0 \end{bmatrix} \underline{a}_s^{(g_i)} \qquad (3.12a)$$

where the following notation is used:

The superscript g_i refers to the i^{th} gyro;

k_{I_i} = the i^{th} gyro's error coefficient due to mass-unbalance along the gyro's input axis, in deg/hr/g

k_{S_i} = the i^{th} gyro's error coefficient due to mass-unbalance along the spin axis, in deg/hr/g

$\underline{a}_s^{(g_i)}$ = the sensed acceleration vector in ft/sec^2, expressed in the coordinates of the i^{th} gyro

$k_{m.u.}$ = $0.15068493 \times 10^{-6}$ is the factor which converts deg/hr/g to rad/sec/(ft/sec^2)

$\epsilon(g_i)$ = drift rate of the i^{th} gyro in rad/sec

The arrangement of the gyro coordinates on the IMU_s is shown in Figure 3.2.

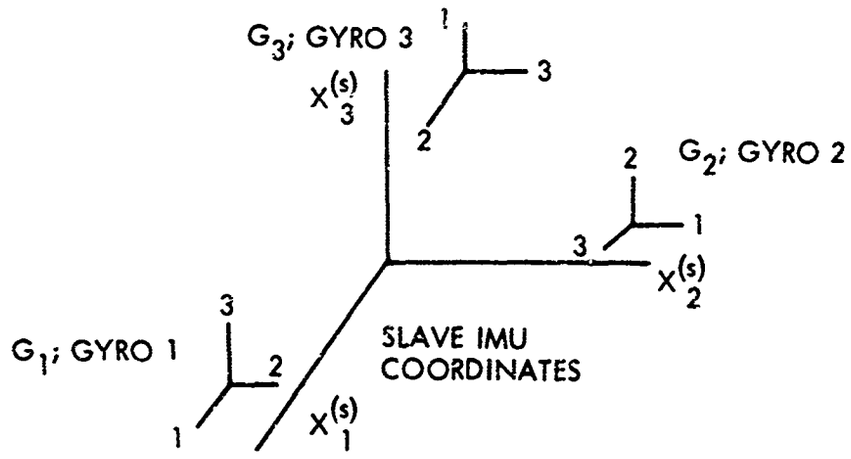


Figure 3.2 Orientation of Slaved IMU Gyros

After introducing the required coordinate transformations, the acceleration error, in slaved IMU coordinates, due to mass unbalance drifts is given by

$$\Delta \underline{a}_{m.u.} = \left(\begin{array}{ccc|ccc} a_{s_1}^{(s)} & 0 & 0 & a_{s_2}^{(s)} & 0 & 0 \\ 0 & a_{s_2}^{(s)} & 0 & 0 & a_{s_3}^{(s)} & 0 \\ 0 & 0 & a_{s_3}^{(s)} & 0 & 0 & a_{s_1}^{(s)} \end{array}\right) \left(\begin{array}{c} k_{I_1} \\ k_{I_2} \\ k_{I_3} \\ \hline k_{S_1} \\ k_{S_2} \\ k_{S_3} \end{array}\right) = \left|a_s^{(s)}\right|_2 \cdot \underline{k}_{m.u.} \qquad (3.12b)$$

The total velocity error due to mass unbalance drifts between samples is then given by:

$$\underline{v}_{m.u.}(k+1,k) = k_{m.u}\int_{k\Delta}^{(k+1)\Delta}\left\{\left[a_s^{(s)}(t)\right]_1\left\{\int_0^t\left[a_s^{(s)}(\tau)\right]_2 d\tau\right\}dt\right\}\underline{k}_{m.u.} \tag{3.12c}$$

The matrix product in Equation (3.12c) is

$$\left[a_s^{(s)}\right]_1 \cdot \left[a_s^{(s)}\right]_2 = \left[\begin{array}{ccc:ccc} 0 & -a_{s_3}a_{s_2} & a_{s_2}a_{s_3} & 0 & -a_{s_3}a_{s_3} & a_{s_2}a_{s_1} \\ a_{s_3}a_{s_1} & 0 & -a_{s_1}a_{s_3} & a_{s_3}a_{s_2} & 0 & -a_{s_1}a_{s_1} \\ a_{s_2}a_{s_1} & a_{s_1}a_{s_2} & 0 & -a_{s_2}a_{s_2} & a_{s_1}a_{s_3} & 0 \end{array}\right] \tag{3.12d}$$

where the superscript s has been dropped. Thus each element in the integrated 3x6 matrix of Equation (3.12c) will involve a term of the form

$$I^{i,j}(k+1,k) \equiv \int_{k\Delta}^{(k+1)\Delta} a_i^{(s)}(t)\left\{\int_0^t a_j(t')dt'\right\}dt \; ; \; i,j = 1,2,3 \tag{3.12e}$$

It is convenient to define

$$\left[v_s^{(s)}\right]_{m.u.} = \left[v_s^{(s)}(k+1,k)\right] \equiv \int_{k\Delta}^{(k+1)\Delta}\left[a_s^{(s)}(t)\right]_1 \cdot\left\{\int_0^t\left[a_s^{(s)}(\tau)\right]_2 d\tau\right\}dt \tag{3.12f}$$

so that Equation (3.12c) may be written as

$$\Delta \underline{v}_{m.u.}(k+1,k) = k_{m.u.}\left[v_s^{(s)}(k+1,k)\right]_{m.u.} \underline{k}_{m.u.} \qquad (3.12g)$$

The estimator equations are then obtained by replacing

$$k_{mis} \cdot \left(v_s^{(s)}\right)_{mis}$$

in Equation (3.6c) by

$$\left[k_{mis}\left(v_s^{(s)}\right)_{mis} \;\vdots\; k_{m.u.}\left(v_s^{(s)}\right)_{m.u.}\right]$$

In the case of this configuration, it is assumed that the random accelerations are uncorrelated, so that $A_{III} = 0$ in Equation (3.6e). Further, the vector b_{III} in Equation (3.6c) is replaced by the 3x3 matrix of Equation (2.5b).

3.5 Initial Misalignment Angles and Constant Gyro Drifts

In this section we obtain the equations for estimating the initial gyro misalignments, Ψ_i, and the constant gyro drift terms ε_i, $i = 1,2,3$. The constant drift of the i^{th} gyro, ε_i, is expressed in deg/hr. The arrangement of the gyros is as shown in Figure 3.2. The velocity error expressed in IMU_S coordinates between samples will thus be

$$\Delta \underline{v}_{d.r.}(k+1,\,k) = k_{d.r.}\int_{k\Delta}^{(k+1)\Delta}\left(a_s^{(s)}(t)\right)\cdot\left\{\int_o^t \underline{\varepsilon}\,dt'\right\}dt$$

$$= k_{d.r.}\int_{k\Delta}^{(k+1)\Delta} t\cdot\left[a_s^{(s)}(t)\right]dt\cdot\underline{\varepsilon} \qquad (3.13a)$$

In this equation $\left[a_s^{(s)}(t)\right]$ is the matrix defined in Chapter II, Section 2.2, and $k_{d.r.} = 0.48481368\times10^{-5}$ is the conversion factor from deg/hr to rad/sec. Next we define

$$\left[v_s^{(s)}\right]_{d.r.} = \int_{k\Delta}^{(k+1)\Delta} t\left[a_s^{(s)}(t)\right]dt \qquad (3.12b)$$

so that

$$\Delta \underline{v}_{d.r.}(k+1,k) = k_{d.r.}\left[v_s^{(s)}\right]_{d.r.}\underline{\varepsilon} \qquad (3.12c)$$

The estimator equations are then obtained by replacing $k_{mis}\left(v_s^{(s)}\right)_{mis}$ in Equation (3.6c) by $\left[k_{mis}\left(v_s^{(s)}\right)_{mis} \vdots k_{d.r.}\left(v_s^{(s)}\right)_{d.r.}\right]$. For this estimator it is assumed that the random accelerations are uncorrelated, so that in Equation (3.6e) $A_{III} = 0$, and b_{III} is replaced by Equation (3.5b).

For convenience in the discussion of the experimental results, the following notation is used for the configurations of the various slaved IMU error models:

Configuration IA: Identification of initial misalignment angles $\underline{\Psi}$, using correlated random acceleration.

Configuration IB: Identification of the initial misalignment angles $\underline{\Psi}$, using uncorrelated random acceleration.

Configuration II: Identification of the initial misalignment angles $\underline{\Psi}$, and the constant gyro drifts $\underline{\epsilon}$.

Configuration III: Identification of the initial misalignment angles $\underline{\Psi}$, and the mass-unbalance gyro drifts, $\underline{k}_{m.u.}$.

CHAPTER IV

EXPERIMENTAL RESULTS

It is quite useful to obtain an idea of the type of results which are possible for the trajectory optimization which is considered in later chapters. In this chapter we give some experimental results concerning the identification of the error parameters for configurations IA, IB, II and III. The effects of state noise, measurement noise, and the nominal trajectory are considered. The nominal trajectory is reasonably assumed to be specified by the thrust acceleration and by the trajectory pitch angle (which is assumed very nearly equal the vehicle's pitch angle). The vehicle considered is assumed to have a fixed acceleration profile, so that the only degree of freedom in specifying the nominal trajectory is its pitch angle, μ. These functions, $\mu(t)$, $t \in [t_o, t_o + T]$, are also constrained in a certain sense. Generally, the pitch angle is constrained to be $\pm$ 10 degrees from some reference value, and the pitch rate might be similarly constrained. Typically, the nominal trajectories would be as shown in Figure 4.1. It should be noted that most of the simplifying assumptions are made in order to facilitate the parametric study, and are not conceptual in nature.

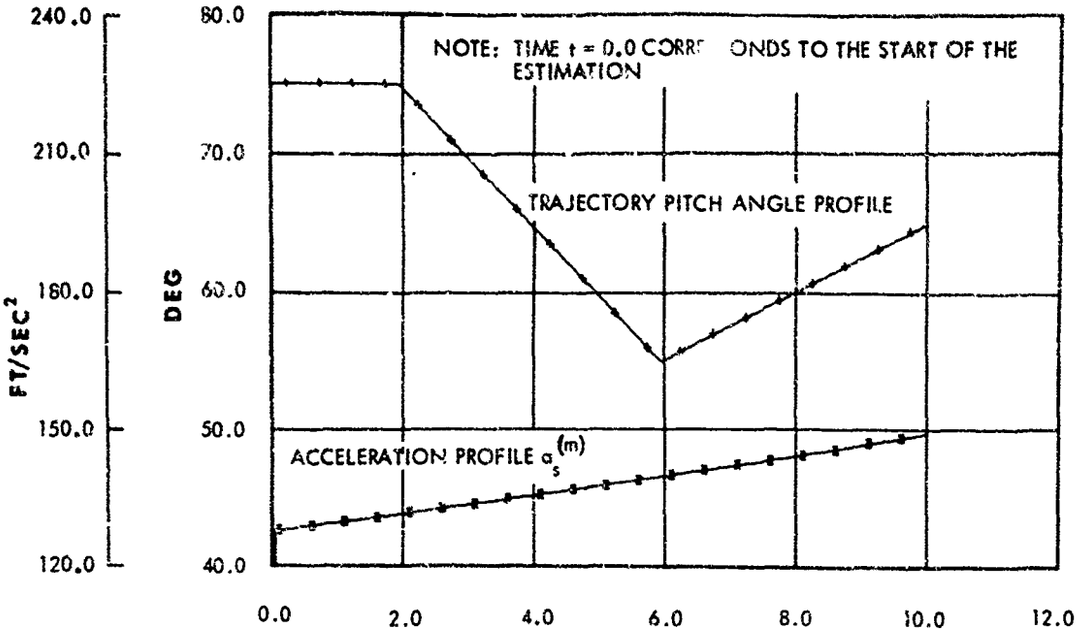


Figure 4.1 Reference Thrust Acceleration Profile and Trajectory Pitch Angle Profile

A discussion of the various initial conditions (a priori estimates, etc.) may be found in [3]. The notation used below is essentially as follows:

$\sigma_n^2 = \phi_o$, the correlated random acceleration power

$\sigma_R^2 \cdot I$ the uncorrelated random acceleration covariance matrix.

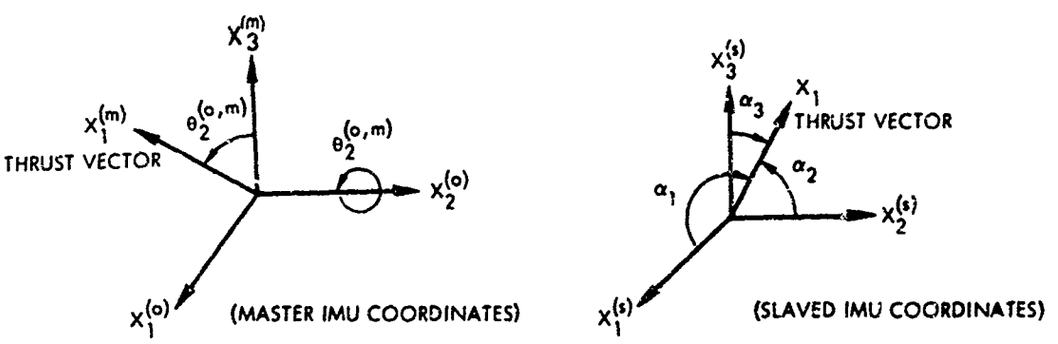
$\sigma_W^2 \cdot I$ the observation noise covariance matrix.

$\sigma_{\hat{\psi}_i}$ the standard deviation in the estimated misalignment angles.

$\sigma_{\hat{\epsilon}_i}$ the standard deviation in the estimated constant gyro drift rates.

$\sigma_{\hat{k}_{S_i}}, \sigma_{\hat{k}_{I_i}}$ the standard deviation in the estimated mass-unbalance gyro drift terms.

The initial IMU_S orientation angles, $\underline{\theta}_0^{(o,s)}$ do not seem to affect $\sum_{i=1}^{3} \sigma_{\hat{\Psi}_i}^2$. However, the $\sigma_{\hat{\Psi}_i}$ of the individual estimated $\Psi_i, \hat{\Psi}_i$, do depend on this orientation. For this reason an orientation in which the accelerometers each sense about the same thrust acceleration was chosen for most of the experimental work to minimize the maximum of the $\sigma_{\hat{\Psi}_i}$. The geometry is as shown below.



The thrust vector in IMU_0 coordinates is

$$\underline{a}_s^{(o)} = a \begin{pmatrix} \sin\theta_2^{(o,m)} \\ 0 \\ \cos\theta_2^{(o,m)} \end{pmatrix} \equiv a \begin{pmatrix} S\phi \\ 0 \\ C\phi \end{pmatrix}$$

and it is necessary to pick the three fixed rotations $\theta_1^{(o,s)}$, $\theta_2^{(o,s)}$, $\theta_3^{(o,s)}$ such that the three direction cosines $\alpha_1, \alpha_2, \alpha_3$ are equal. Thus

$$\underline{a}_s^{(s)} = \frac{a}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = T\left(\theta_3^{(o,s)}, \theta_2^{(o,s)}, \theta_1^{(o,s)}\right) \underline{a}_s^{(o)}$$

If $\theta_2^{(o,m)} = (TP - 90^\circ) = -25^\circ$, and $\theta_2^{(o,s)} = TP - 45^\circ$, the solution to the above equation gives $\theta_1^{(o,s)} = 48.9^\circ$, $\theta_2^{(o,s)} = 20^\circ$, and $\theta_3^{(o,s)} = -22.5^\circ$.

4.1 Experimental Results for Correlated Random Acceleration (Configuration IA)

The correlated random acceleration model is studied in considerable detail in [3] For this configuration the gyro drift rates are assumed negligible so that only the initial misalignment angles are estimated. The effects of the sampling rate (data processing rate), Δ, are shown in Figure 4.2. The sampling time is varied between 0.025 sec and 0.100 sec to determine how much improvement in the identification of the Ψ_i and the correlated noise components is possible. The most suitable sampling rate depends on the particular correlated noise model, in that it should be fast enough to obtain a reasonable representation of the correlated random acceleration. The effects of the random acceleration rms power levels and measurement noise levels are shown in Figure 4.3. It seems that it is of advantage to structure the estimator for correlated random accelerations if it is only required to estimate the initial misalignment angles and if the correlated noise is of a sufficiently low "frequency" (say less than three cycles/sec). It is also of advantage to consider the correlated

model if it is of prime importance to estimate the random motion with all error parameters known, as, for example, in launching from a mother ship. If errors in addition to the ψ_i are also to be estimated it does not seem possible to use the time correlation in the random acceleration to advantage.

Since the results for the correlated noise model depend on many parameters and on the specific noise model considered, further discussion will not be made here. Considerable experimental results for the correlated random acceleration model may be

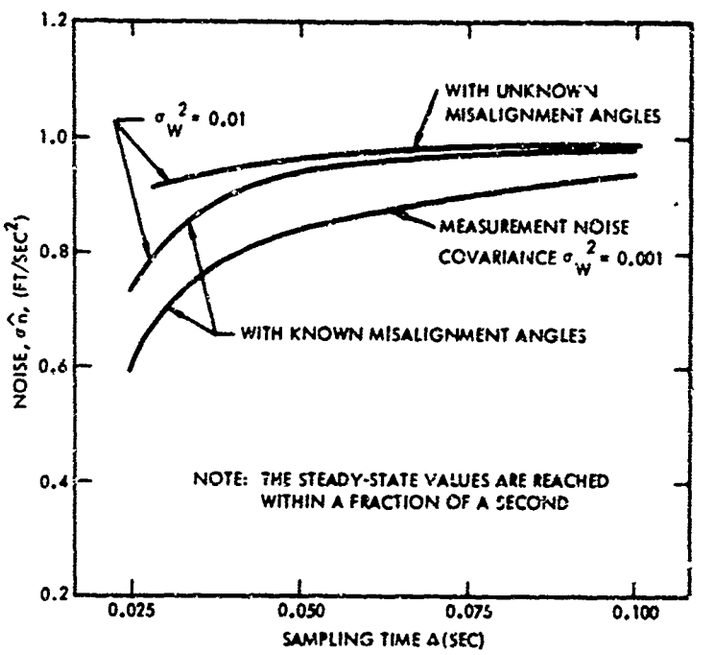


Figure 4.2 Steady-State Standard Deviation of the Estimate of the Acceleration Noise, $\sigma_{\hat{n}}$ for Various Sampling Times Δ

found in [3]. In this case a parametric study is not as readily made, or as meaningful, as for the uncorrelated noise model which is considered next. Since one of the main objects of this study is to consider the maneuver which allows that the identification of the errors be accomplished in an optimal fashion (Chapter VI), the uncorrelated model is used because it provides a better setting for this purpose.

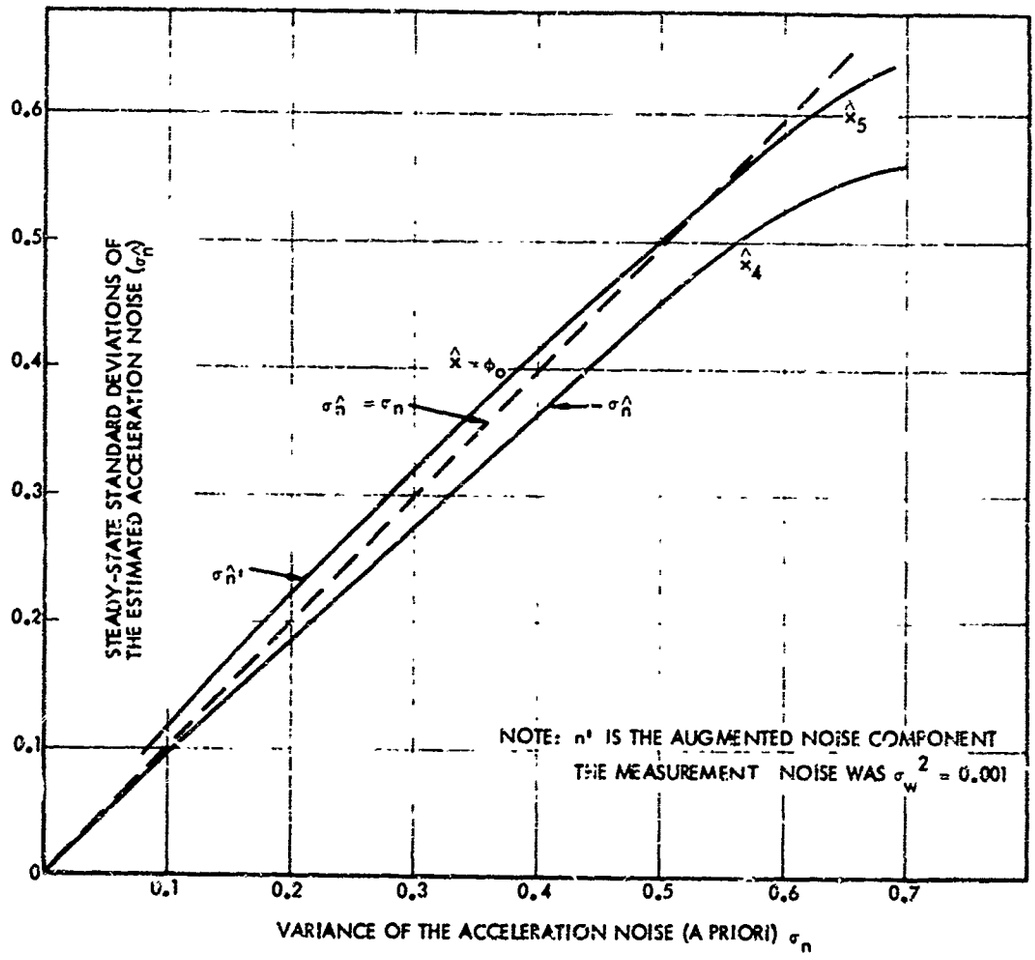


Figure 4.3a Steady-State Standard Deviations for the Estimated Random Acceleration $\hat{n}$

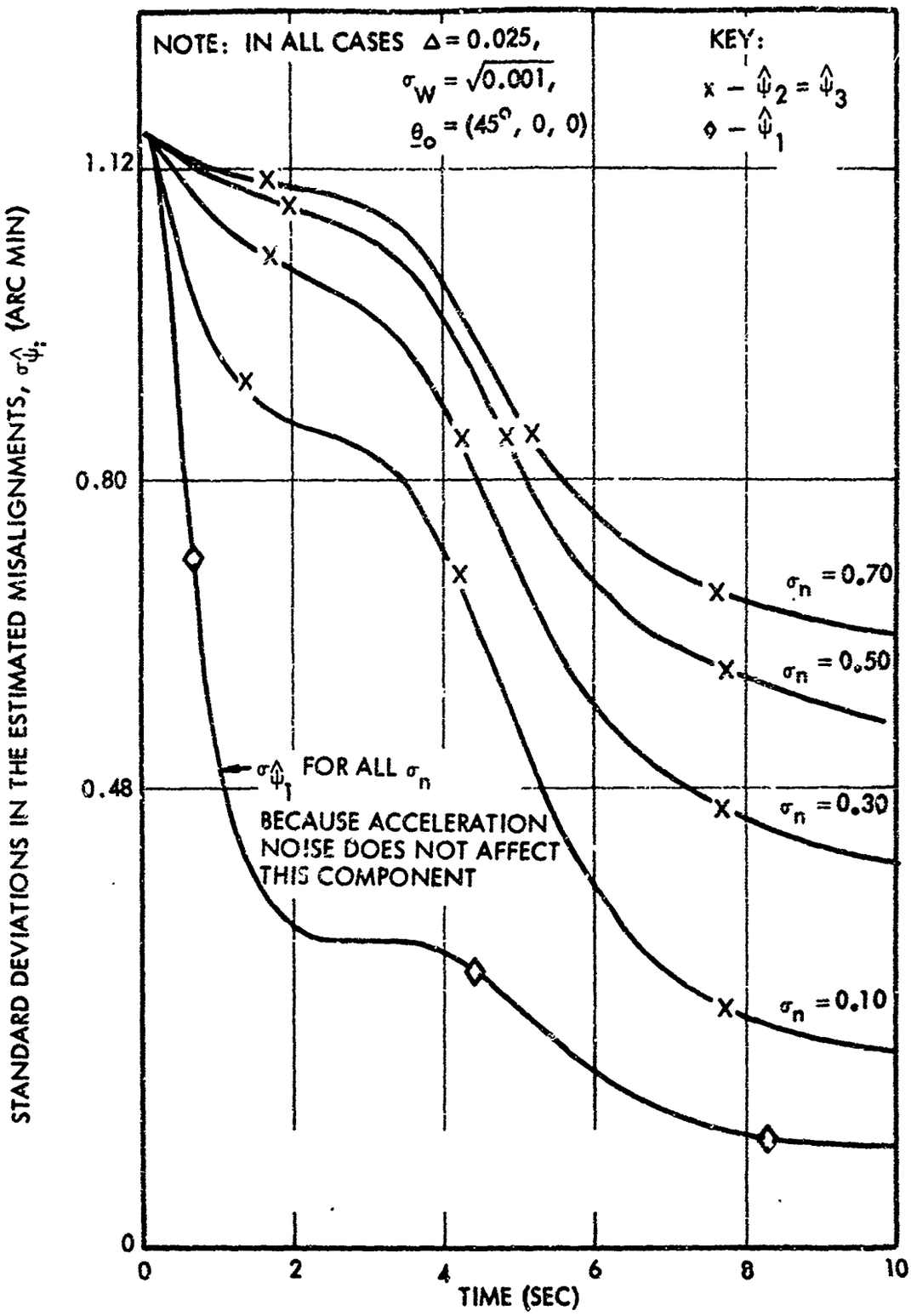


Figure 4.3b $\sigma_{\hat{\Psi}_1}$ for various σ_n using the Nominal shown in Fig 4.1

4.2 Experimental Results for Initial Misalignment Angles and Uncorrelated State Noise (Configuration IB)

In this case the data the data processing rate was fixed at $\Delta = 0.10$ seconds and the "equal acceleration" orientation discussed above was used. The effect of the state noise, measurement noise, and trajectory pitch profile are considered. Figure 4.5 illustrates the sigma ψ_i for different values of measurement noise standard deviation, σ_W. The random acceleration variance σ_R, for these cases was fixed at $\sigma_R = 10^{-3}$, and Nominal Trajectory IB-7, Figure 4.4, was used. Figure 4.6 illustrates the sigma ψ_i for different values of random acceleration variance, σ_R^2. A pertinent discussion of the effects of process noise (random acceleration) and measurement noise on the covariance equation P, is given in Section 4.5.

The manner in which the IMU_S is aligned relative to IMU_O (through the angles $\underline{\theta}_O^{(O,S)}$) and the trajectory pitch profile enter the variance equation are outlined in Chapter V. Since the object is to minimize the trace of this equation, the angles $\underline{\theta}_O^{(O,S)}$ and μ could be considered as control variables in this minimization. Theoretical considerations regarding the existence of a minimum with respect to $\underline{\theta}_O^{(O,S)}$ and μ are made in Chapter VI. Only experimental results are presented here which might give some insight into this optimization problem. The sigma ψ_i for the

IMU_S initial misalignment angles, $\underline{\theta}_o^{(o,s)}$, listed below were considered first.

$$\underline{\theta}_o^{(o,s)} = (48.9, 20, -22.5)^*, \ (0,0,45)^*, \ (0,45,0)^*, \ (45,0,0)^*$$

To make a comparison $\hat{\sigma} \equiv [\sum_{i=1}^{3} \sigma_{\hat{\psi}_i}^2]^{1/2}$ was plotted for the four different initial IMU_S alignment angles, $\underline{\theta}_o^{(o,s)}$. In the four cases presented, $\hat{\sigma}$ did not depend on the initial alignment angles, $\underline{\theta}_o^{(o,s)}$, indicating that a minimum of the sum of the variances of the $\sigma_{\hat{\psi}_i}^2$ does not exist as a function of $\underline{\theta}_o^{(o,s)}$. It was clear from the results that the individual $\sigma_{\hat{\psi}_i}$ do depend on $\underline{\theta}_o^{(o,s)}$, so that more realistic performance criteria might be to minimize the maximum $\sigma_{\hat{\psi}_i}$ as a function of $\underline{\theta}_o^{(o,s)}$. Theoretical difficulties will arise with such a criterion because the maximum principle cannot be used to obtain necessary conditions for optimality. However, it does seem that the equal acceleration orientation satisfies the above criterion. For the remainder of this study, the IMU_S orientation is such that $\underline{\theta}_o^{(o,s)} = (48.9 \ , \ 20 \ , \ -22.5 \ )^*$ (equal acceleration orientation) was chosen because the three $\sigma_{\hat{\psi}_i}$ are about the same, and there is no reason to have them otherwise at this point. It should be noted that the random acceleration and the measurement noise were assumed to be such that the variances on each component were equal, i.e., each component of velocity difference. In the case in which these

variances are different for each component of velocity difference, the initial IMU orientation, $\underline{\theta}_o^{(o,s)}$ would surely affect $\hat{\theta}$.

The sigma $\hat{\Psi}_i$ for four different μ profiles are shown in Figure 4.7. The four different nominal trajectories used are shown in Figure 4.4, and are essentially such that μ approaches its lower value of 55 deg stepwise, but at different rates. In all cases the steady-state $\sigma_{\hat{\Psi}_i}$ for each constant portion of μ (for each of the four trajectories) is the same. Only the rate at which these steady-state values of $\sigma_{\hat{\Psi}_i}$ are approached is different. If larger σ_R and σ_W were used, it is possible that the steady-state values would not be reached, and that the $\sigma_{\hat{\Psi}_i}$ time histories for the four nominal trajectories would be different. Actually, the size of the step determines how much larger the error due to Ψ_i in the velocity differences, $\underline{v}^{(o,s)}$, is than that due to random acceleration and measurement noise. The larger this difference, the more confidence the estimator has in choosing the $\hat{\Psi}_i$. The corresponding minimum variance estimates, $\hat{\Psi}_i$, for two cases are shown in Figure 4.8. Further discussion of the nominal trajectories are given in the next section.

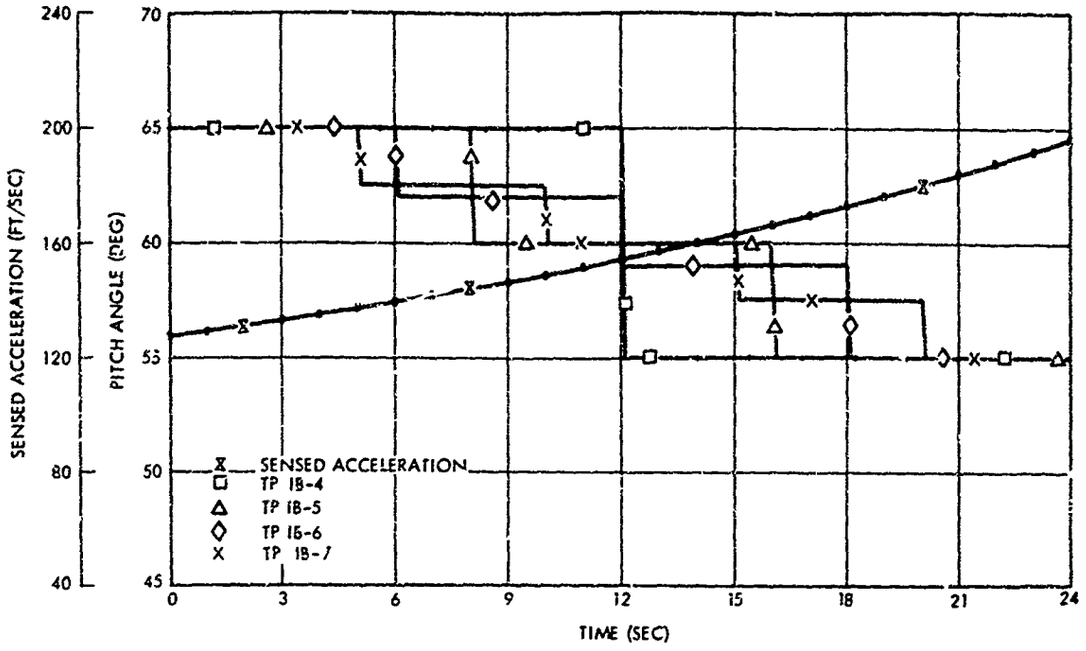


Figure 4.4 Nominal Trajectories IB-4 through IB-7

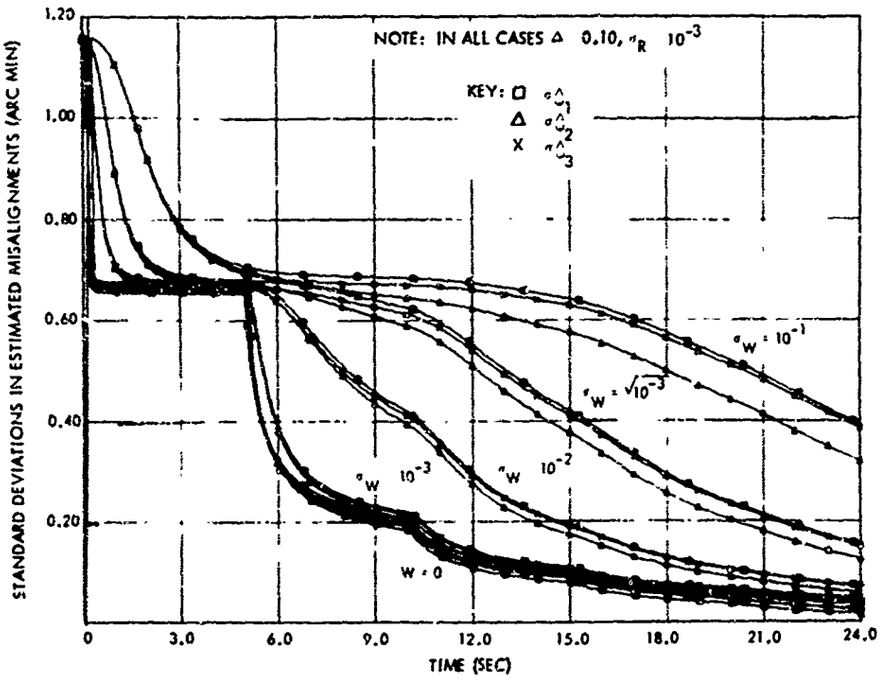


Figure 4.5 Effect of σ_W on the Estimation of the ψ_i Using Nominal IB-7

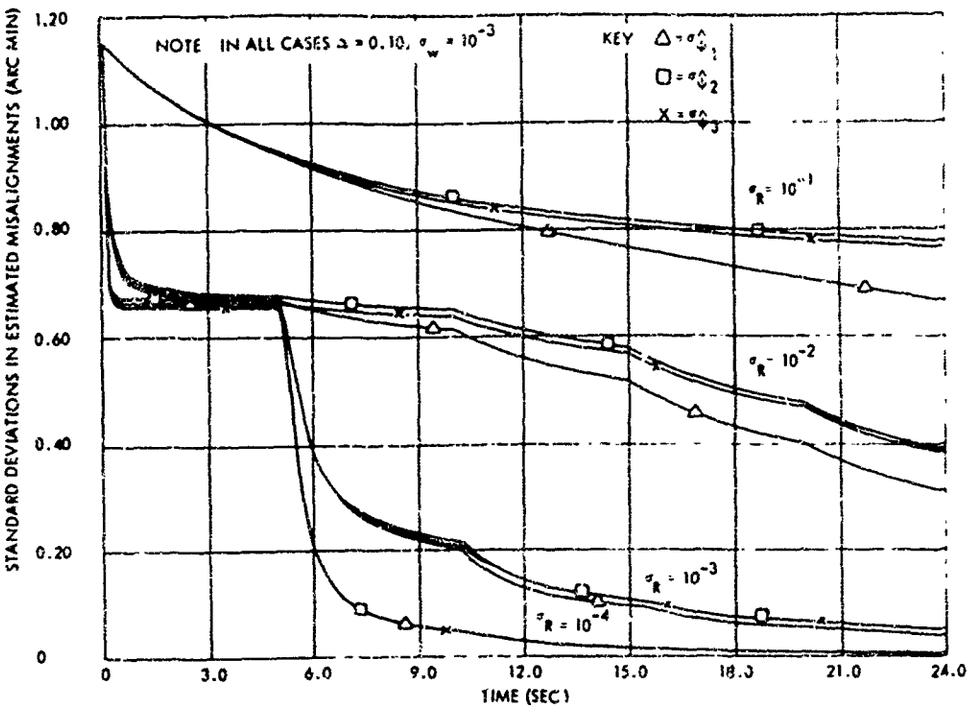


Fig 4.6 Effect of σ_R on the Estimation of the ψ_i, using Nominal IB-7

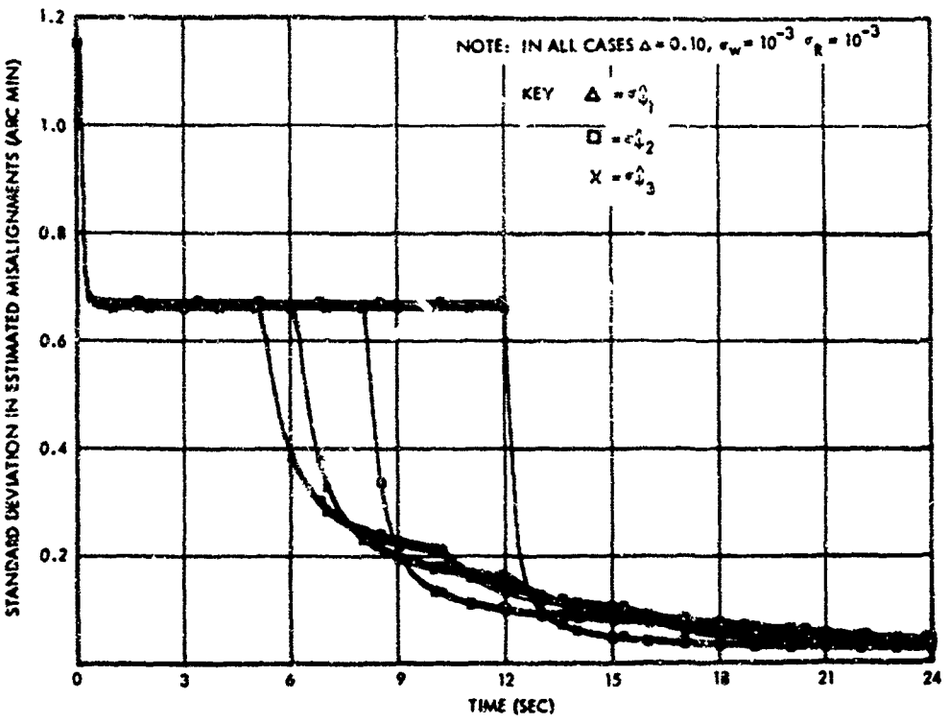
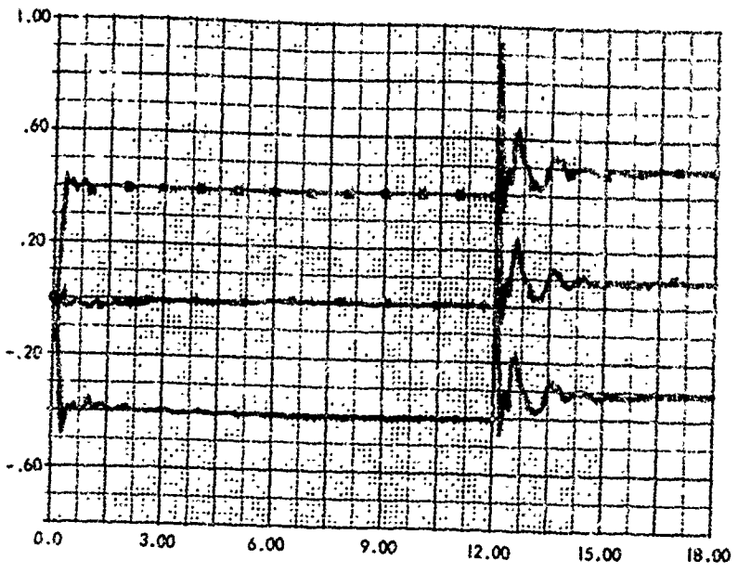
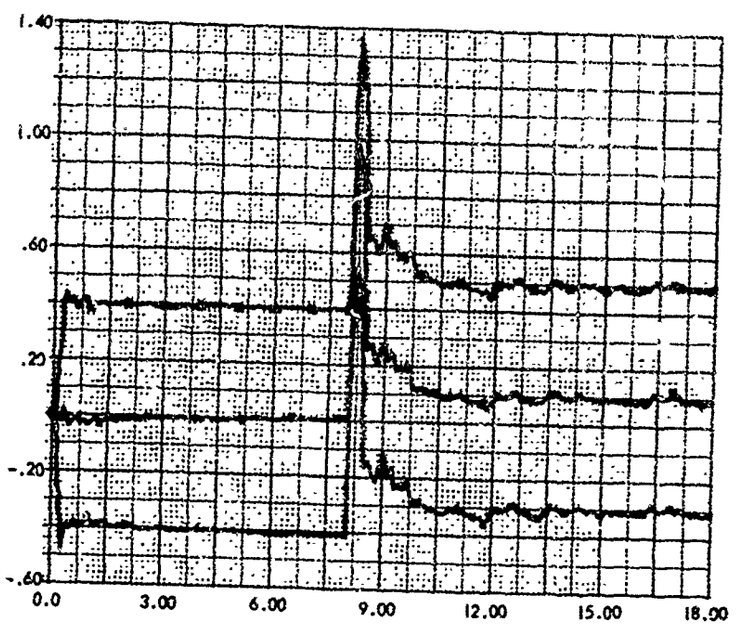


Fig. 4.7 Effect of the Pitch Profile (Fig. 4.4) on the Estimation of the ψ_i



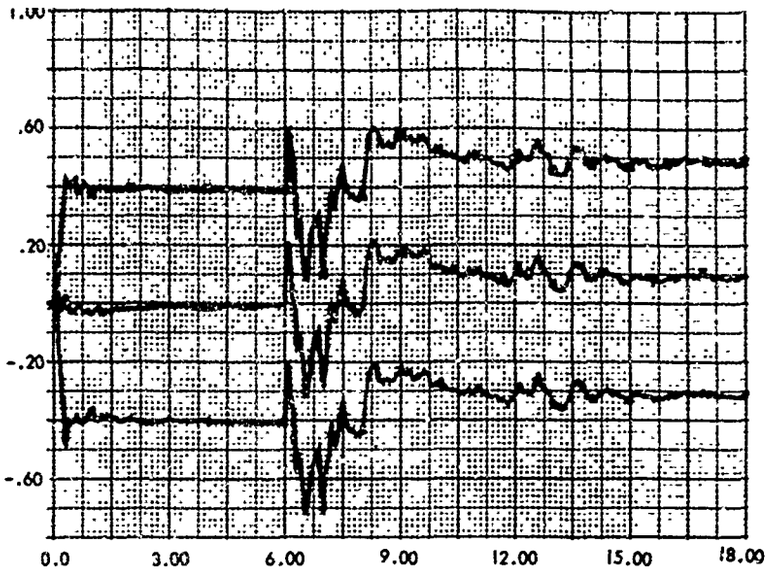
a) Nominal Trajectory IB-3



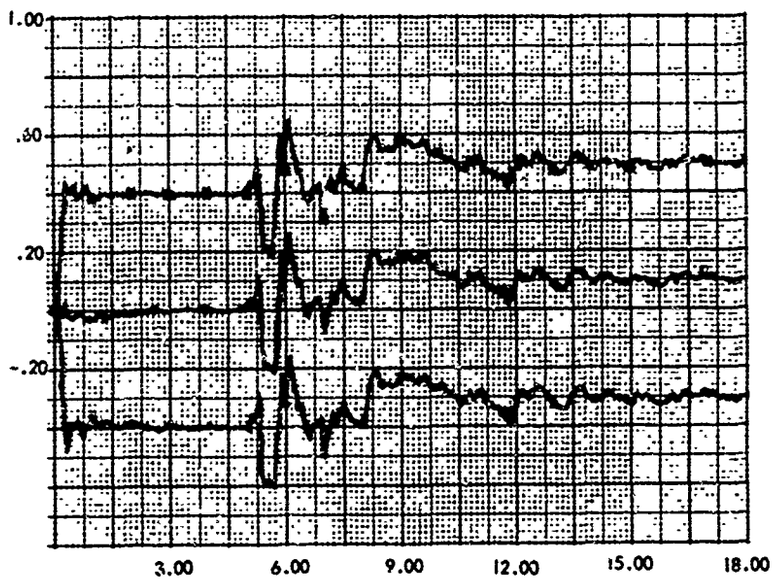
b) Nominal Trajectory IB-4

Note: In both cases $\Delta = .10$, $\sigma_W = \sigma_R = 10^{-3}$, $\underline{\psi} = (.5, .1, -.3)^*$

Figure 4.8a Estimates of ψ_i, Using Nominals IB-3 and IB-4



c) Nominal Trajectory IB-5



d) Nominal Trajectory IB-6

Note: In both cases $\Delta = .10$, $\sigma_W = \sigma_R = 10^{-3}$, $\underline{\psi} = (.5, .1, -.3)^*$

Figure 4.8b Estimates of ψ_i Using Nominals IB-5 and IB-6

4.3 Experimental Results for Initial Misalignment Angles and Constant Gyro Drift Rates (Configuration II)

Effect of Random Acceleration and Measurement Noise on Configuration II

Figures 4.9 and 4.10 show the effects that the measurement noise, σ_W, and the random acceleration, σ_R, have on the estimates of the constant gyro drifts, ϵ_i, and the initial misalignment angles, Ψ_i. The value of σ_W does not affect the $\hat{\Psi}_i$ too much (Figure 4.9a), except $\sigma_{\hat{\Psi}_i}$ decreases faster for smaller σ_W. The steady-state values of the $\sigma_{\hat{\Psi}_i}$ are independent of σ_W, as is discussed in Section 4.5. The effect of σ_W on the $\sigma_{\hat{\epsilon}_i}$ is stronger than for the $\sigma_{\hat{\Psi}_i}$, especially between $\sigma_W = 10^{-3}$ and 10^{-4} ft/sec (Figure 4.9b). No conclusions about the steady-state values of the $\sigma_{\hat{\epsilon}_i}$ as a function of σ_W can be deduced, since the $\sigma_{\hat{\epsilon}_i}$ are still decreasing at t = 24 sec.

The effect of the random acceleration (σ_R) on the estimation of the Ψ_i and ϵ_i is illustrated in Figures 4.10a and 4.10b respectively. For $\sigma_R = 10^{-3}$ ft/sec^2, there is an appreciable steady-state value (~0.05 arc min) in the $\sigma_{\hat{\Psi}_i}$, that is, appreciable when compared to the error due to the constant gyro drifts, ϵ_i. As shown in Figure 4.10b, the effect of the random acceleration noise levels on the

estimation of the ϵ_i is considerably more pronounced. For $\sigma_R = 10^{-5}$ and 10^{-6} ft/sec^2, the $\sigma_{\hat{\epsilon}_i}$ are about the same values, decreasing significantly to about 0.0125 deg/hr at t = 24.0 sec. These results indicate that it would be desirable to keep the measurement noise levels such that $10^{-5} < \sigma_W \leq 10^{-4}$ ft/sec and the random acceleration noise levels such that $10^{-5} < \sigma_R \leq 10^{-4}$ ft/sec^2.

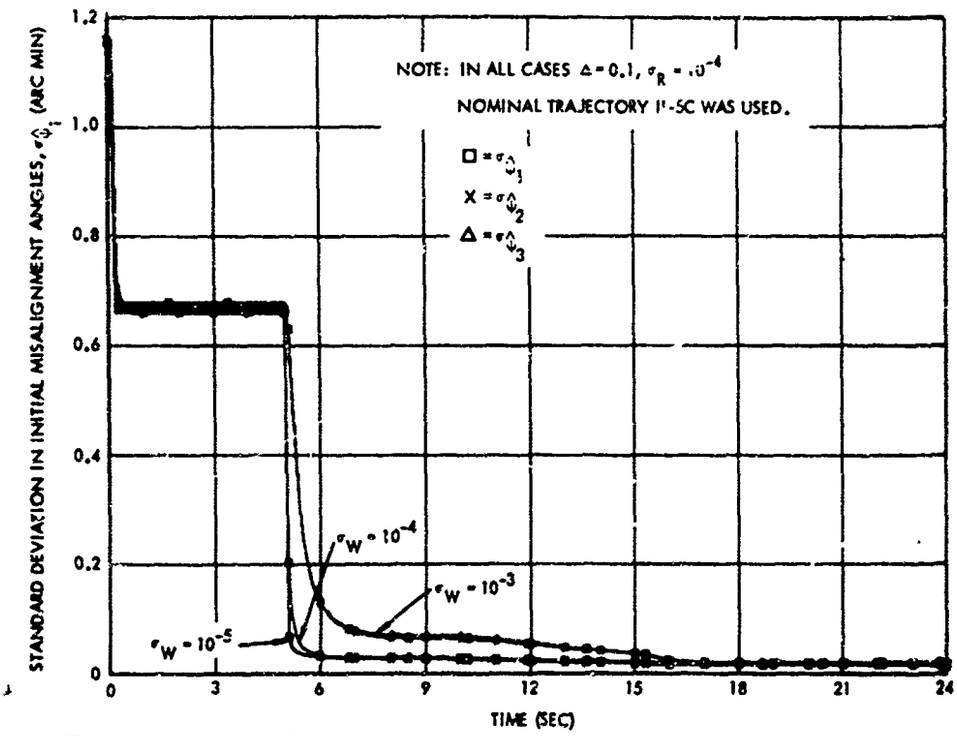


Figure 4.9a Effects of Measurement Noise, σ_W, on the Estimates of the Initial Misalignment Angles

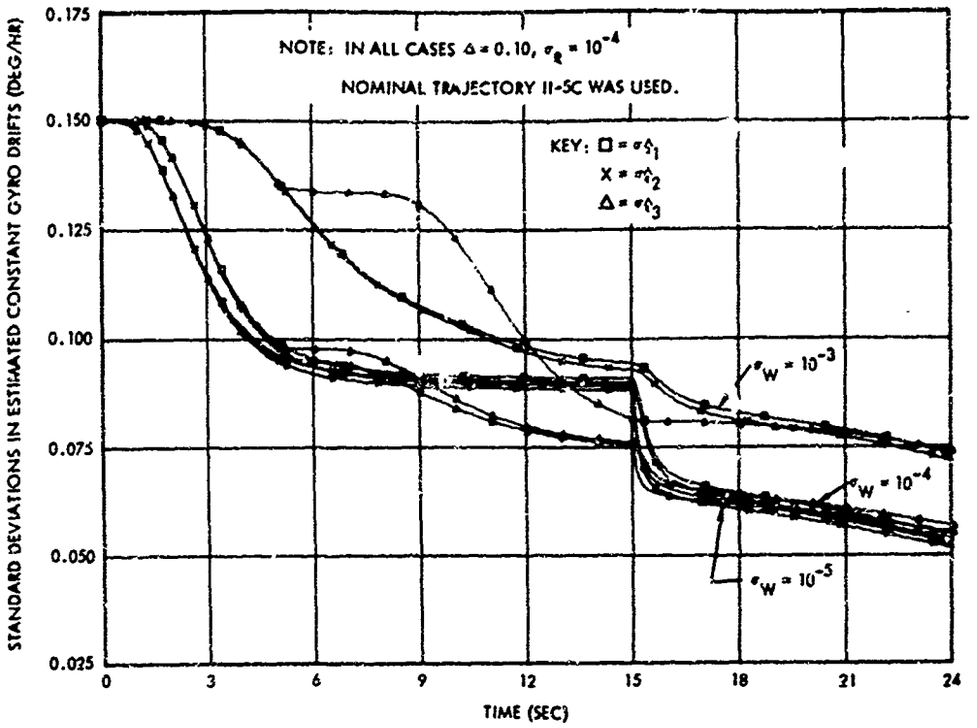


Figure 4.9b Effects of Measurement Noise, σ_W, on the Estimation of the Constant Gyro Drift Rates

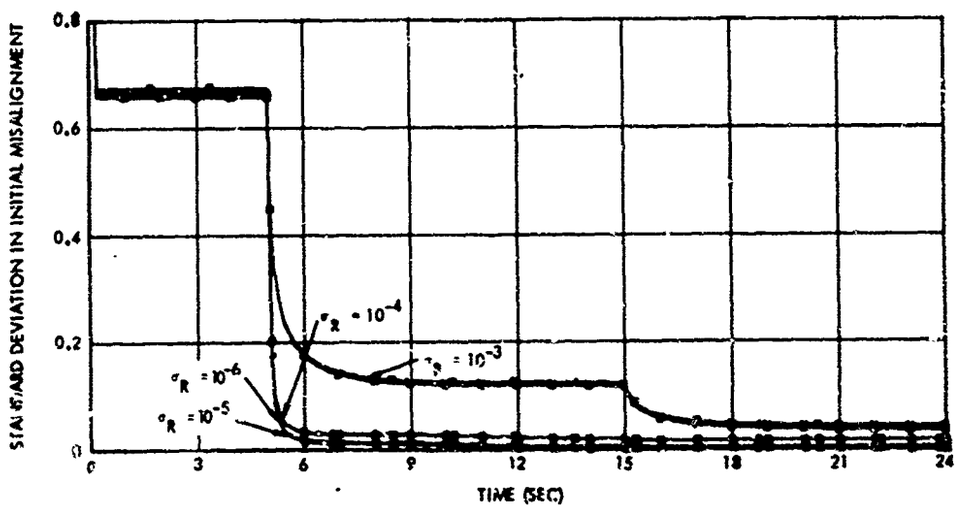


Figure 4.10a Effect of the Random Acceleration, σ_R, on the Estimates of the Initial Misalignment Angles

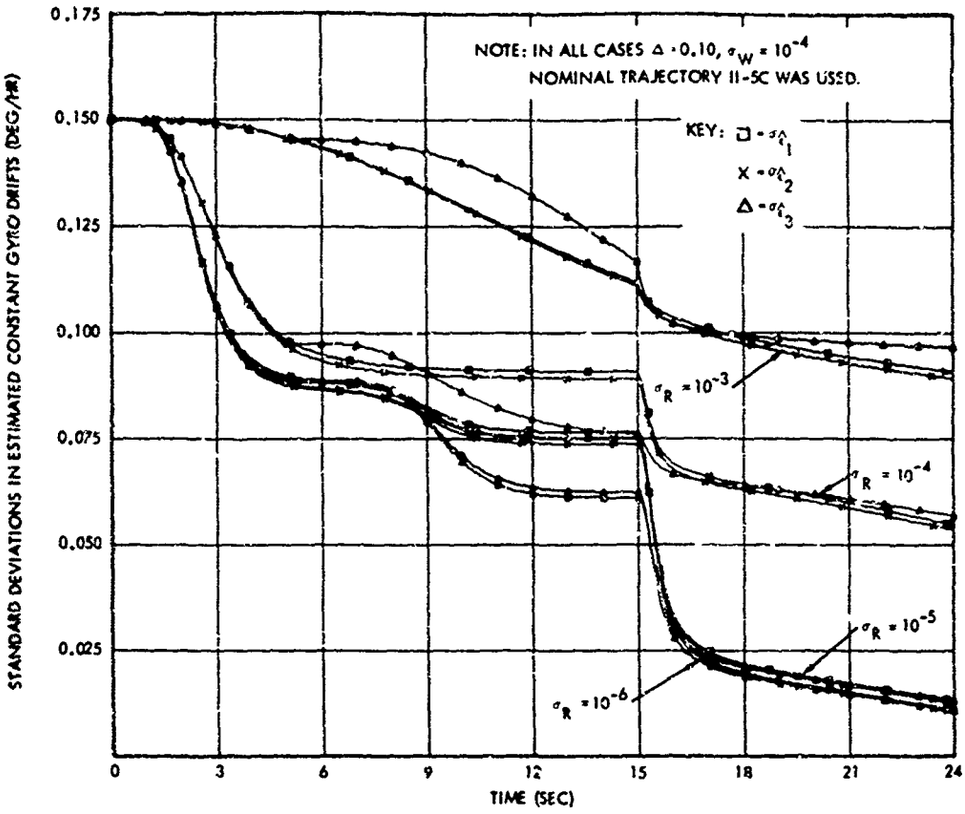


Figure 4.10b Effect of the Random Acceleration, σ_R, on the Estimates of the Constant Gyro Drifts

Effect of the Nominal Trajectory on Configuration II

The trajectory pitch profile, μ, is varied to study its effect on the estimation. The variations are such that the effects of the rate of change of μ, total amount of change in μ, and direction of the change in μ on the estimates can be observed. It should be noted that the constant drift rates, ϵ_i, do not appreciably affect the estimates of the initial misalignment angles, Ψ_i,

because the errors due to ϵ_i are much smaller than those due to the Ψ_i for short estimation times. For this reason the conclusions regarding the Ψ_i deduced from this section can be used in the previous section.

Figure 4.11 shows six trajectory pitch profiles in which they change as ramp functions. The corresponding standard deviations in the $\hat{\Psi}_i$, $\sigma_{\hat{\Psi}_i}$, and the standard deviations in the $\hat{\epsilon}_i$, $\sigma_{\hat{\epsilon}_i}$, are shown in Figures 4.12a and 4.12b respectively. It is seen from Figures 4.12a and 4.12b that the higher rate of changes in TP result in faster decreasing $\sigma_{\hat{\Psi}_i}$ and $\sigma_{\hat{\epsilon}_i}$. However, the steady values in the $\sigma_{\hat{\Psi}_i}$, that is at $t = 24.0$ sec, are about the same (about 0.025 arc min). The $\sigma_{\hat{\epsilon}_i}$ are still rapidly decreasing at $t = 24.0$ and the rate of decrease of the $\sigma_{\hat{\epsilon}_i}$ is proportional to the rate of change of the pitch angle. For TP-2f, the $\sigma_{\hat{\epsilon}_i}$, have decreased from 0.15 deg/hr to about 0.040 deg/hr.

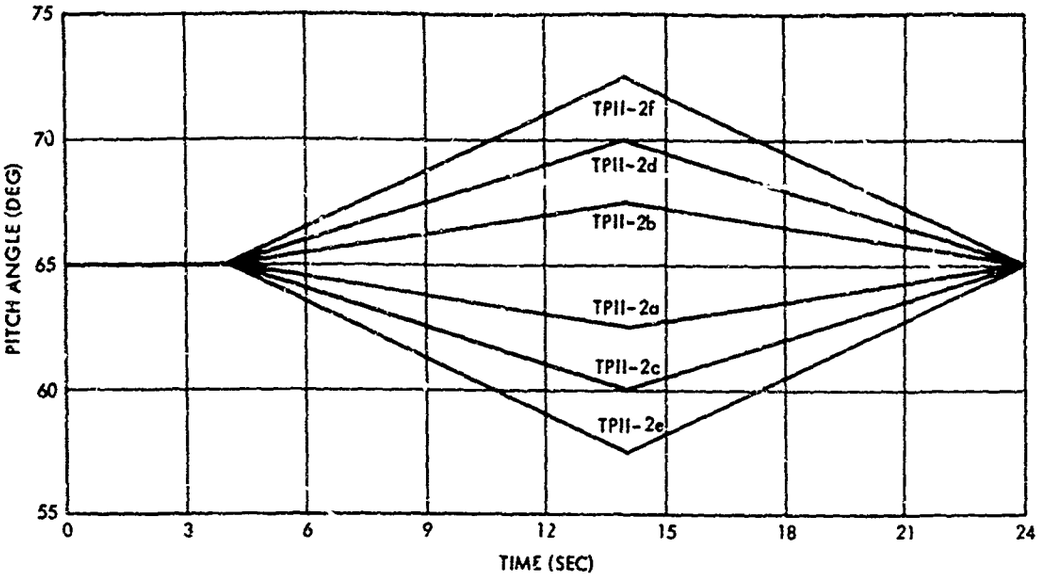


Figure 4.11 Trajectory Pitch Angle Profiles for Nominal Trajectories II-2a through II-2f

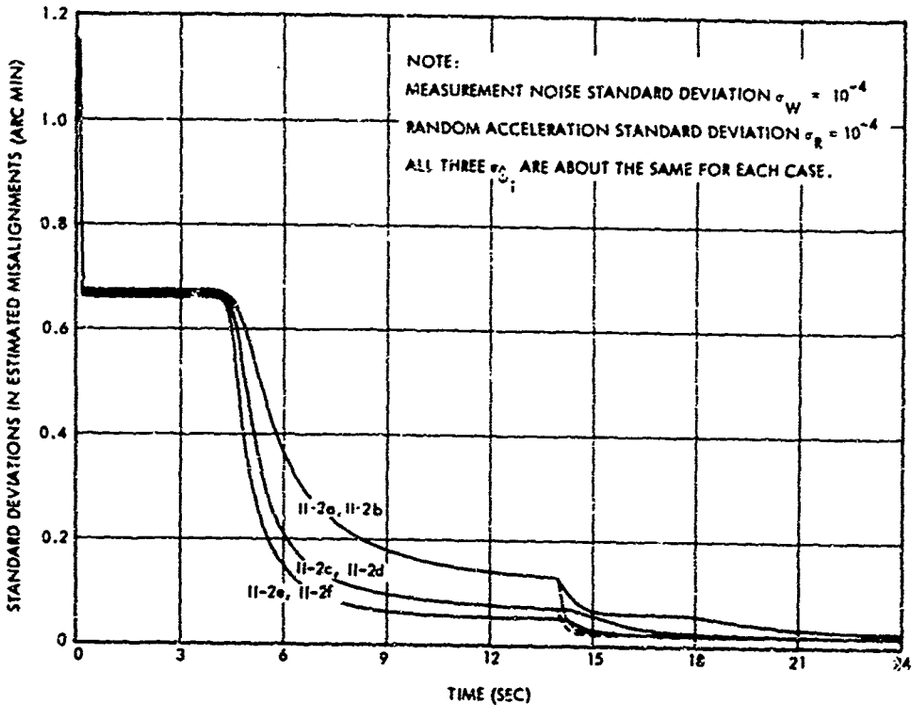


Figure 4.12a Standard Deviations in the Estimated Misalignments, $\sigma_{\hat{\psi}_i}$, for Nominal Trajectories II-2a through II-2f

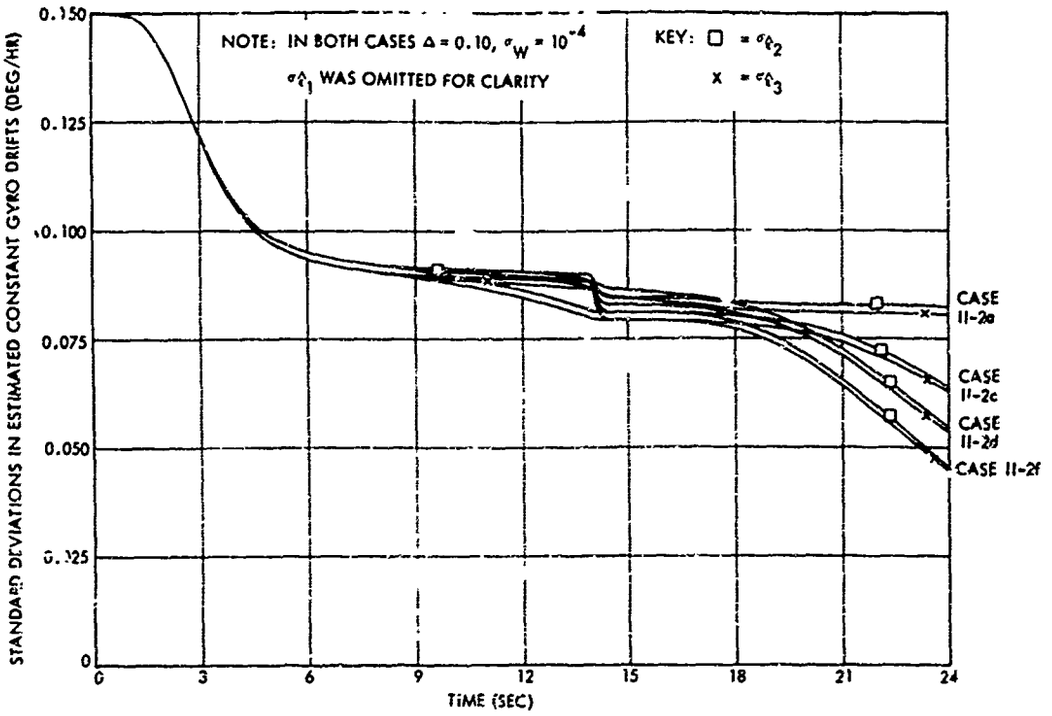


Figure 4.12b Standard Deviations in the Estimated Constant Gyro Drifts, $\sigma_{\hat{\epsilon}_i}$, for Nominal Trajectories II-2a to II-2f

Next, the effect of the total change in TP on the estimates is considered. The pitch profile is changed discretely at t = 10 sec, as shown in Figure 4.13a. These nominal trajectories are referred to as Cases II-3a to II-3g. The corresponding $\sigma_{\hat{\Psi}_i}$ and $\sigma_{\hat{\epsilon}_i}$ are shown in Figures 4.13b and 4.13c, d respectively. The rate of decrease of the sigmas for $\hat{\Psi}_i$ and $\hat{\epsilon}_i$ for the various "step" pitch profiles is proportional to the larger changes in TP. This is reasonable because the larger the change in the pitch profile, the larger the contribution of the Ψ_i and ϵ_i to the

velocity differences, $\underline{v}^{(o,s)}$, as compared to the contribution due to the random acceleration and measurement noises. From Figure 4.14a it is noticed that $\sigma_{\hat{\epsilon}_1}$ and $\sigma_{\hat{\epsilon}_2}$ improve when TP increases (II-3e and II-3g) and it is relatively insensitive to downward changes. On the other hand, Figure 4.14b shows that $\sigma_{\hat{\epsilon}_3}$ improves as TP decreases (II-3d and II-3f) and does not change much when TP increases (II-3g and II-3e). This indicates that a combination of an upward and a downward maneuver would be better for extracting the constant gyro drift ϵ_i. This is further indicated by the results shown in figure 4.15b using nominal 11-5c.

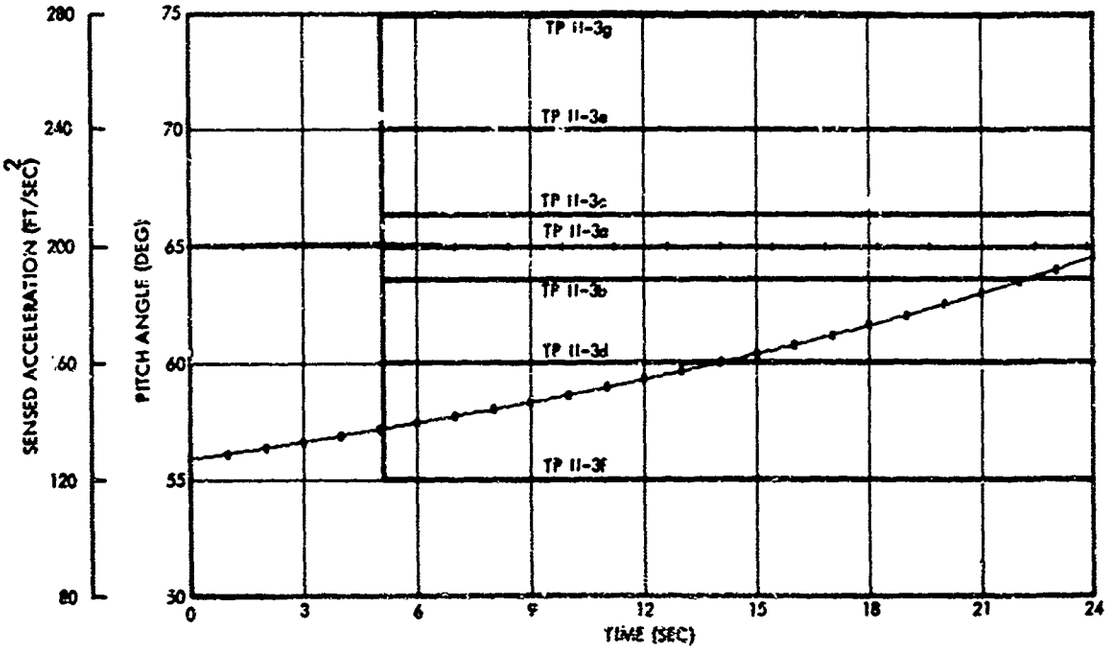


Figure 4.13a Nominal Trajectories II-3a through II-3g

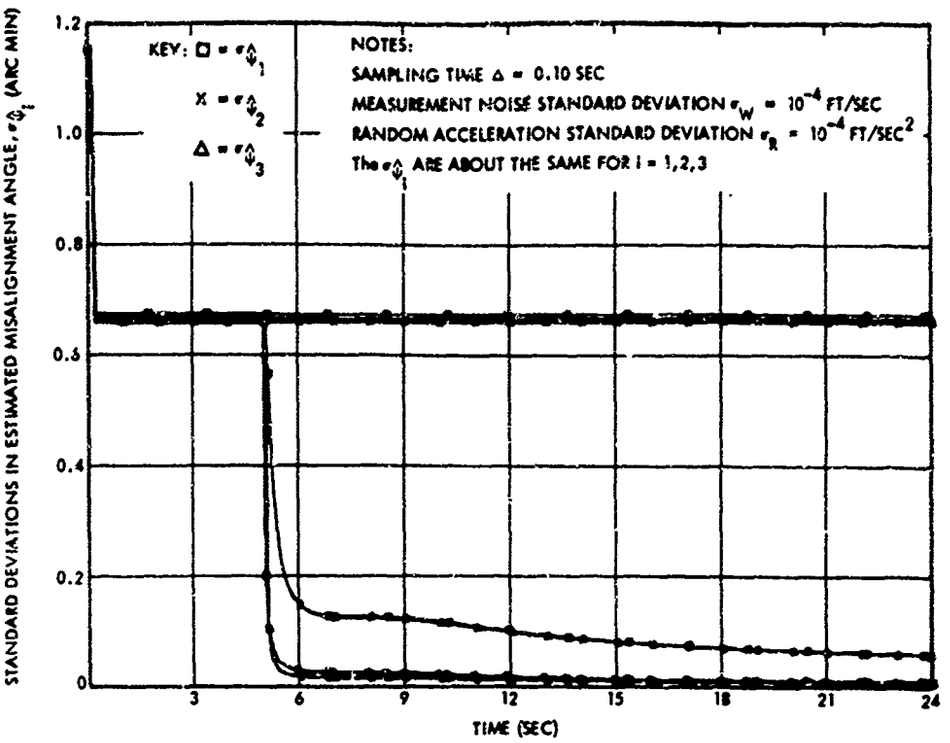


Figure 4.13b $\sigma_{\hat{\psi}_j}$ for Nominals II-3a through II-3g

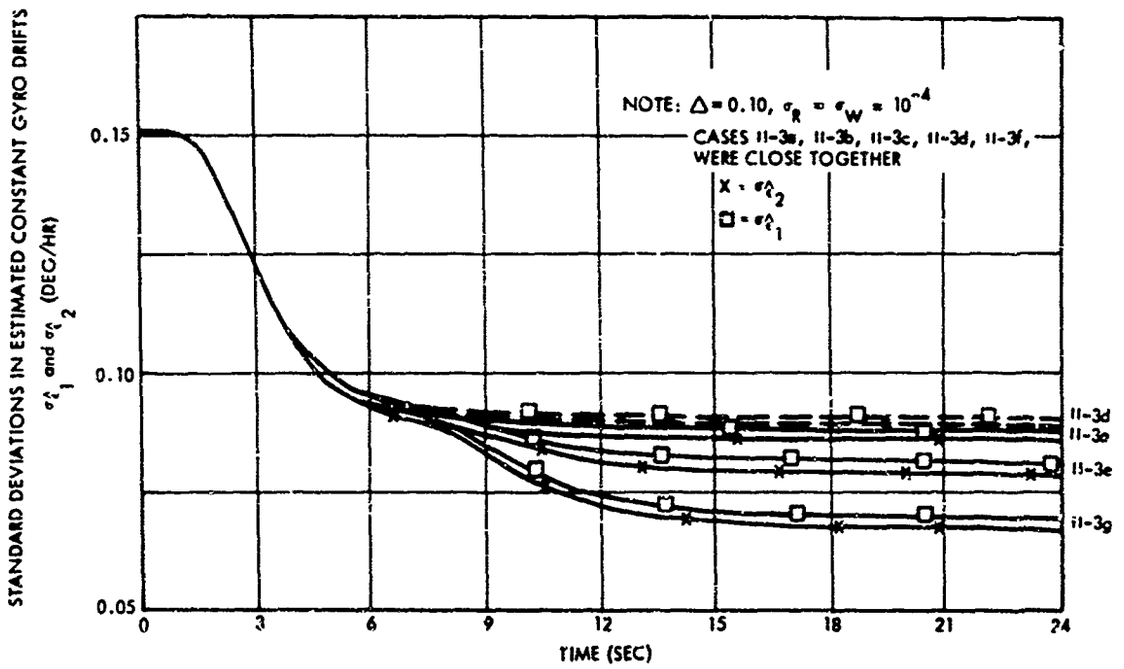


Figure 4.13c $\sigma_{\hat{\epsilon}_1}$ and $\sigma_{\hat{\epsilon}_2}$ for Nominals II-3a through II-3g

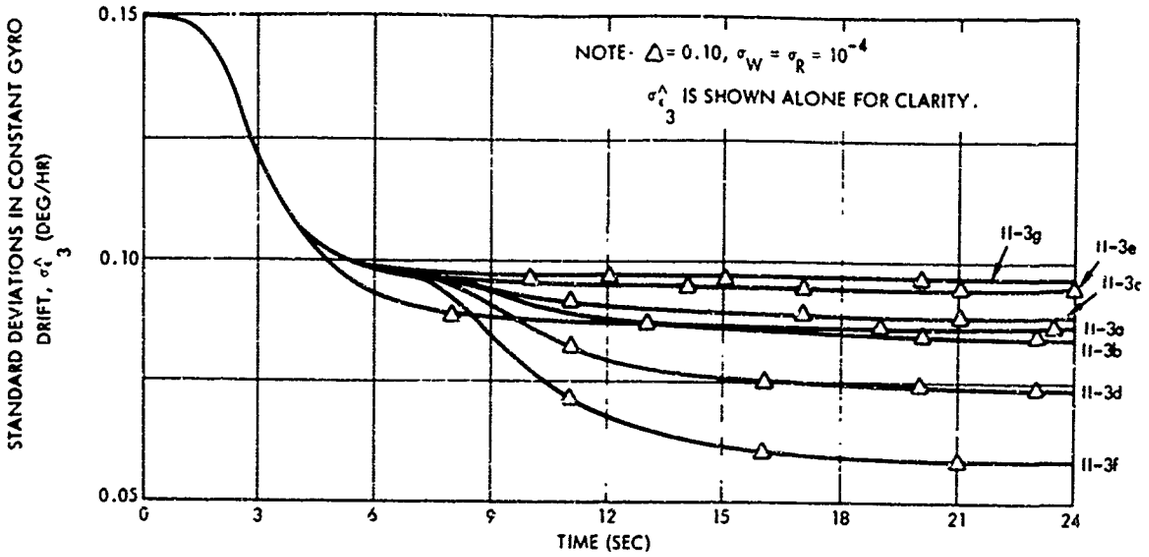


Figure 4.13d Standard Deviations in the Estimated Constant Gyro Drifts, $\sigma_{\hat{\epsilon}_3}$, for Nominals II-3a Through II-3g

The trajectory is next varied in such a way that the step of 10 deg (trajectory TP II-3f) is approached at different rates. This type of variation gives the tradeoff in performance of Configuration II as a function as less abrupt changes in the pitch angle. The trajectory pitch angles are shown in Figure 4.14a. The corresponding $\sigma_{\hat{\psi}_i}$ and $\sigma_{\hat{\epsilon}_i}$ are shown in Figures 4.14b, 4.14c and 4.14d, respectively. As would be expected, the trajectory profile with the most abrupt changes produces better results, for ψ_i as the $\sigma_{\hat{\psi}_i}$ steady out within one second and the most abrupt changes in the trajectory contribute a larger amount to $\underline{v}^{(o,s)}$ for a longer period of time. For the ϵ_i the slower ramps produce better results because the $\sigma_{\hat{\epsilon}_i}$ take longer to steady out, and better results would be obtained if the changes are made later.

It should be noted that if the estimation time is long enough, it doesn't seem to matter how the trajectory changes - as long as it does change. This is true for the $\hat{\Psi}_i$, at least, because the $\sigma_{\hat{\Psi}_i}$ all seem to steady out to the same values. This could be explained in terms of σ_W and σ_R and the discussion of Section 4.5 Note that the slower the ramp decreases, the better the ϵ_i are estimated (contrary to the situation for the Ψ_i). This is clear from Figure 4.14c and 4.14a. Also, $\hat{\epsilon}_3$ is better than $\hat{\epsilon}_1$ and $\hat{\epsilon}_2$ because the trajectory pitch angle is decreased from 65 deg. This behavior was observed for the step profile of Nominal Trajectories II-3a to II-3g.

Figures 4.13 and 4.14 indicate that $\sigma_{\hat{\epsilon}_1}$ is better than $\sigma_{\hat{\epsilon}_2}$ and $\sigma_{\hat{\epsilon}_3}$ if μ changes downwardly from 65 deg, and vice-versa. Figure 4.15b shows the $\sigma_{\hat{\epsilon}_i}$ for a combination of a downward and an upward maneuver. The specific trajectories used are II-5a, b, and c shown in Figure 4.15a with $\sigma_W = \sigma_R = 10^{-4}$. It is clear from Figure 4.15b that the ϵ_i can be equally well estimated, and $\sigma_{\hat{\epsilon}_i}$ is less than 0.055 deg/hr after 24 sec of estimation. If the trajectory pitch profile were varied in an optimal manner, better estimates would surely result.

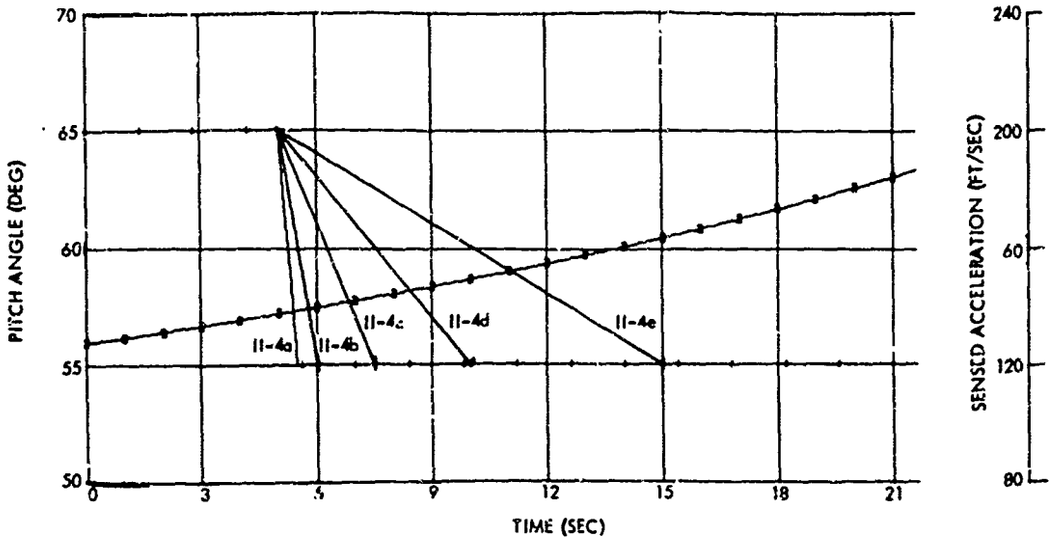


Figure 4.14a Nominal Trajectories II-4a Through II-4e

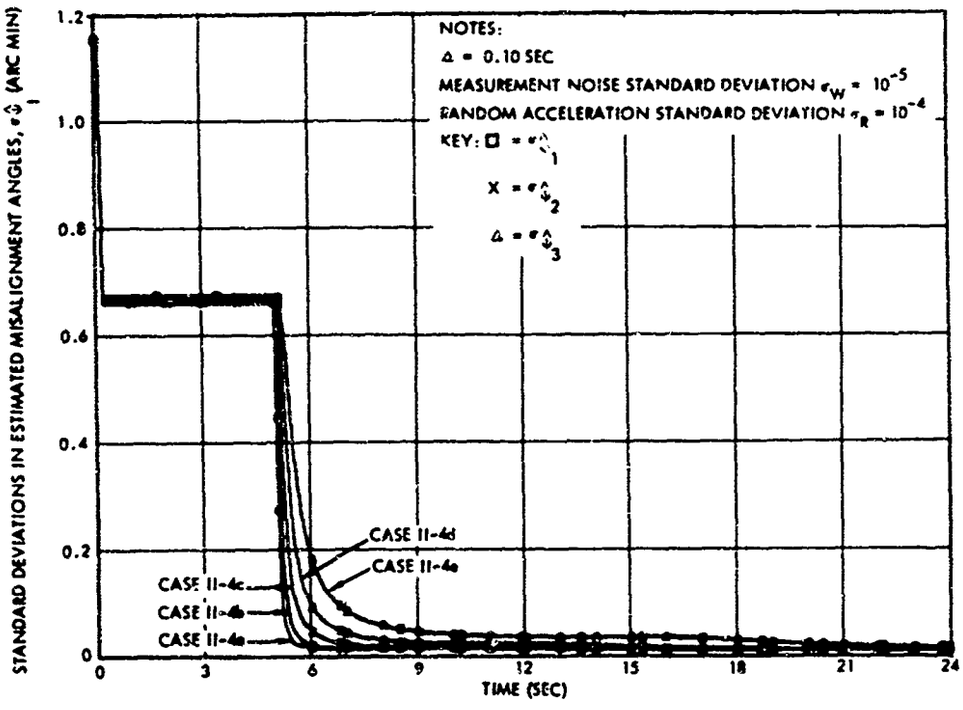


Figure 4.14b Standard Deviations in the Estimated Misalignment Angles, $\sigma_{\hat{\psi}_i}$, for the Various Ramps of Nominal Trajectories II-4a Through II-4e

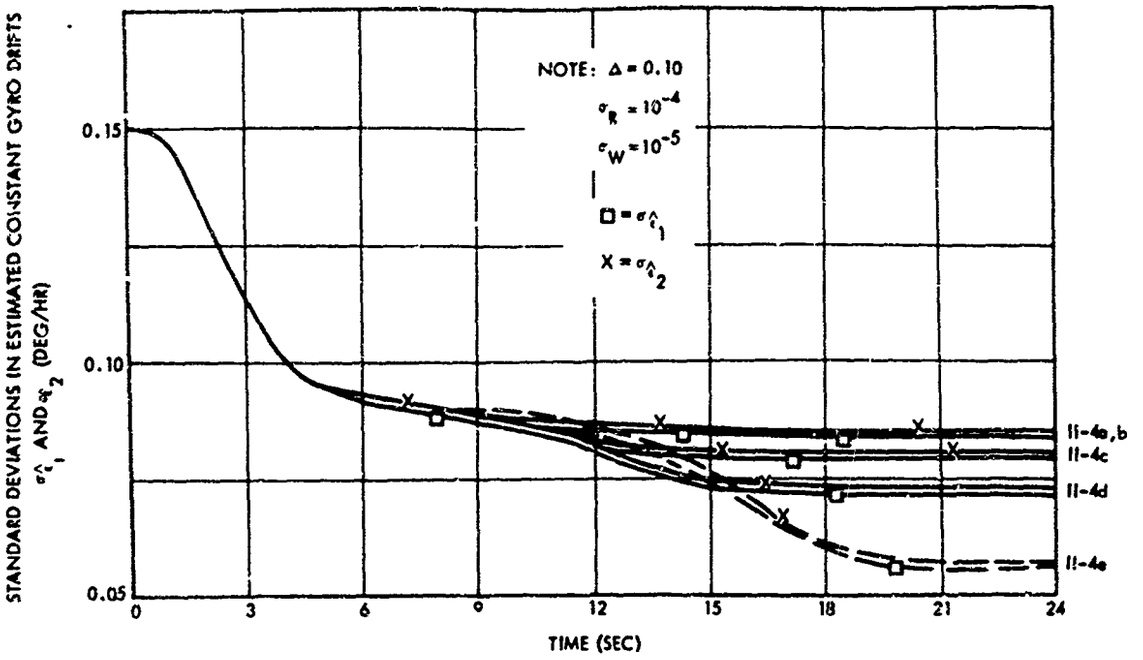


Figure 4.14c Standard Deviations in the Estimated Constant Gyro Drifts, $\sigma_{\hat{\epsilon}_1}$ and $\sigma_{\hat{\epsilon}_2}$, for the Step Profiles

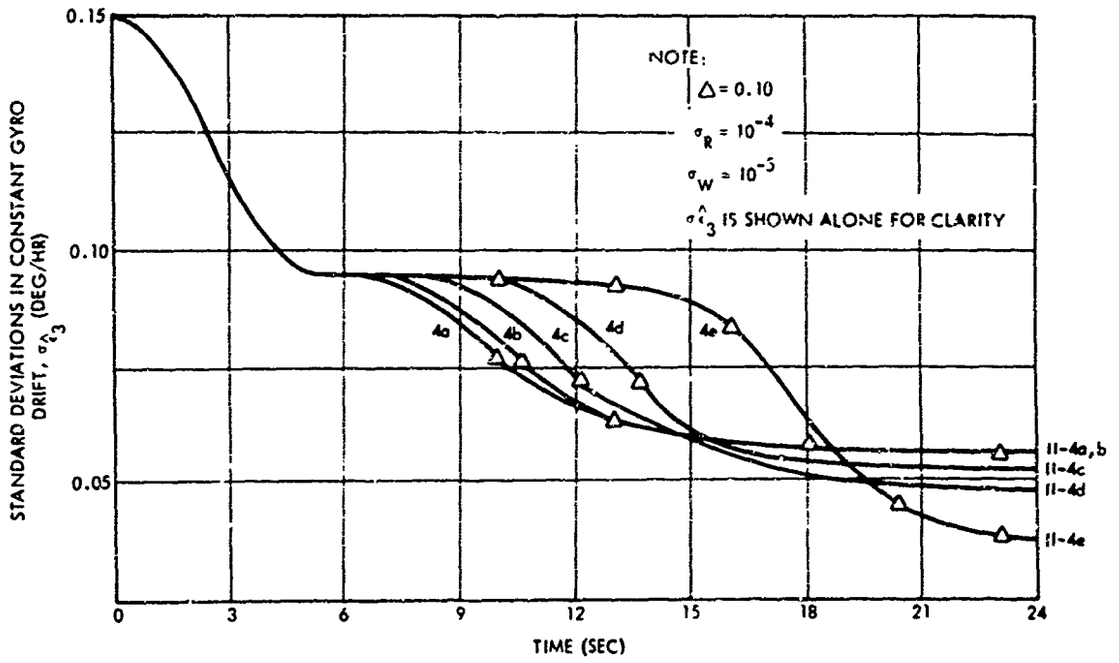


Figure 4.14d Standard Deviations in the Estimated Constant Gyro Drifts, $\sigma_{\hat{\epsilon}_3}$, for the Step Profiles (Nominal Trajectories II-4e)

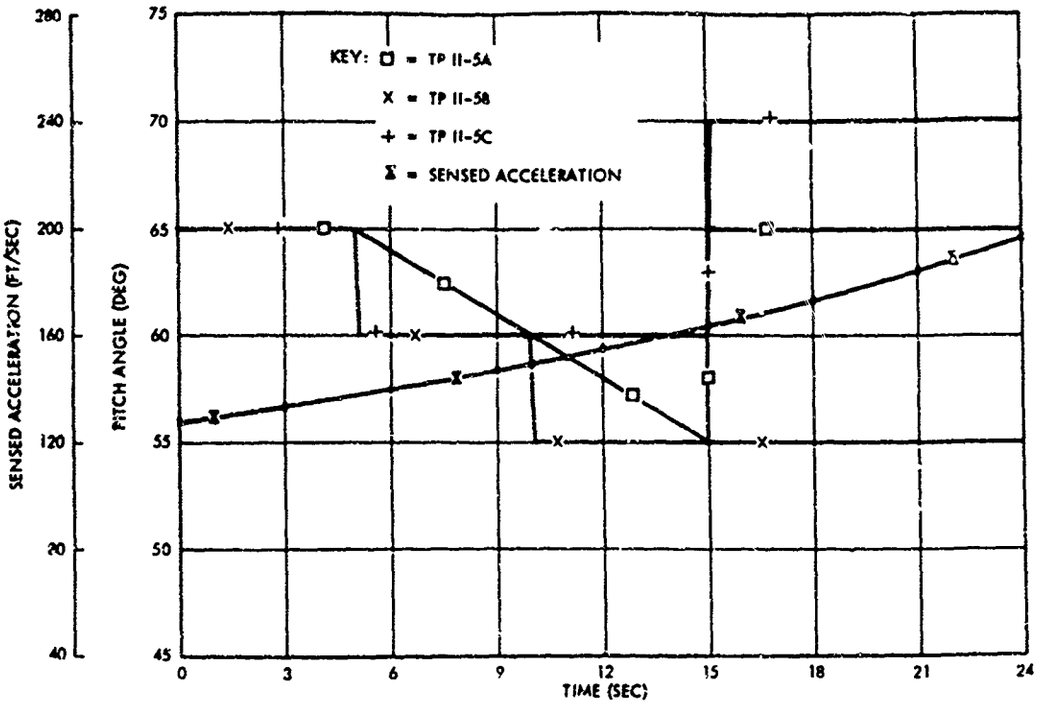


Figure 4.15a Nominal Trajectories II-5a, II-5b, and II-5c

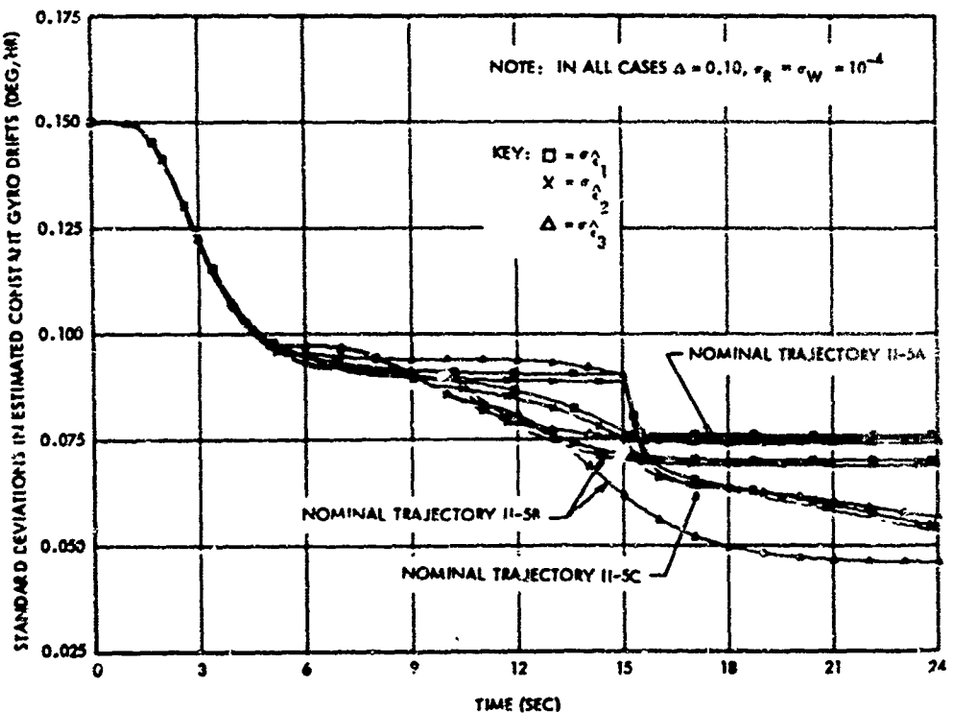


Figure 4.15b Comparison of the Standard Deviations in the Estimated Constant Gyro Drifts for Three Different Nominals of figure 4.15a

4.4 Experimental Results for Initial Misalignment Angles and Mass Unbalance Drift Rates (Configuration III)

The effect of the nominal trajectory and the initial orientation of IMU_S on the estimates has been discussed in Sections 4.2 and 4.3. These effects on Configuration III are not studied here, but it is expected that the results obtained in the previous sections would generally apply, at least for the estimation of Ψ_i. The estimation of the k_{S_i} and k_{I_i} errors would possibly require more trajectory maneuvering.

The first set of standard deviation curves (Figures 4.17 and 4.19) show the effects of the random acceleration on the estimates. The sigma $\hat{\Psi}_i$ are shown in Figure 4.17, and it is clear from these curves that the trajectory pitch angle could have been changed earlier, since the steady-state values of about 0.67 are reached rather quickly. Estimates of the Ψ_i such that $\sigma_{\hat{\Psi}_i} < 0.1$ arc min are obtained in all cases, except for $\sigma_R = 10^{-3}$ ft/sec^2. The spin-axis mass-unbalance standard deviations, $\sigma_{\hat{k}_{S_i}}$, for the different σ_R are shown in Figure 4.19a. Except for the case where $\sigma_R = 10^{-3}$, μ could also have been changed earlier. The input-axis mass-unbalance standard deviations, $\sigma_{\hat{k}_{I_i}}$, are shown in Figure 4.19. These estimates are somewhat better than those for the spin-axis terms. Again the pitch angle could have

been changed earlier, say at t = 5.0 sec . These results show that if $\sigma_W \leq 10^{-4}$ ft/sec and $\sigma_R \leq 10^{-4}$ ft/sec (except the k_{S_i} terms need $\sigma_R < 10^{-4}$) and estimating times are in excess of 20 sec, excellent estimates of Ψ_i, k_{S_i}, and k_{I_i} would be obtained. Based on these results it is conjectured that a well chosen maneuver would improve the estimates, especially the estimates of the spin-axis terms, k_{S_i}.

Next the effect of the measurement noise on the estimates is considered. The nominal trajectory is TP-III-3 in Figure 4.16. This is essentially the same as Nominal Trajectory III-2, except that TP is changed from 65 deg at 5 sec instead of at 10 sec. The corresponding estimation sigmas are shown in Figures 4.18 and 4.20. Typical minimum variance estimates (for $\sigma_W = 10^{-5}$ and $\sigma_R = 10^{-4}$) are shown in Figure 4.21. Generally, the measurement noise does not affect the estimates as in the case of the random acceleration σ_R. Figures 4.18 and 4.20 show that there is not much difference in the sigmas for $\sigma_W = 10^{-3}$ and $\sigma_W = 10^{-4}$.

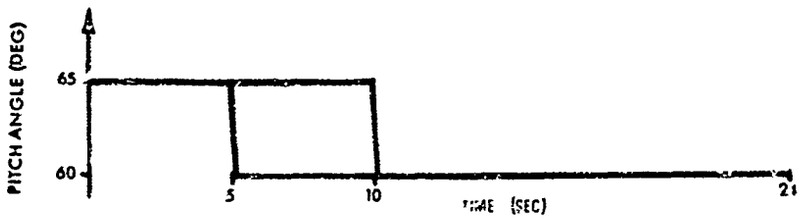


Figure 4.16 Nominal Trajectories III-2 and TP III-3

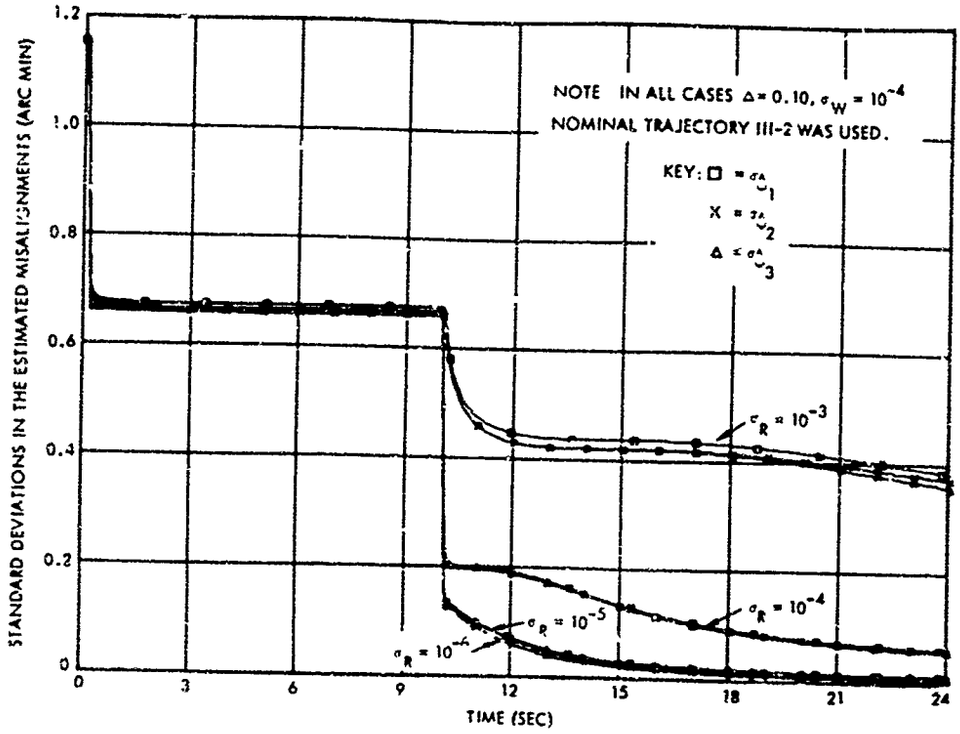


Figure 4.17 Effect of Random Acceleration on the Estimates of Initial Misalignments, ψ_i

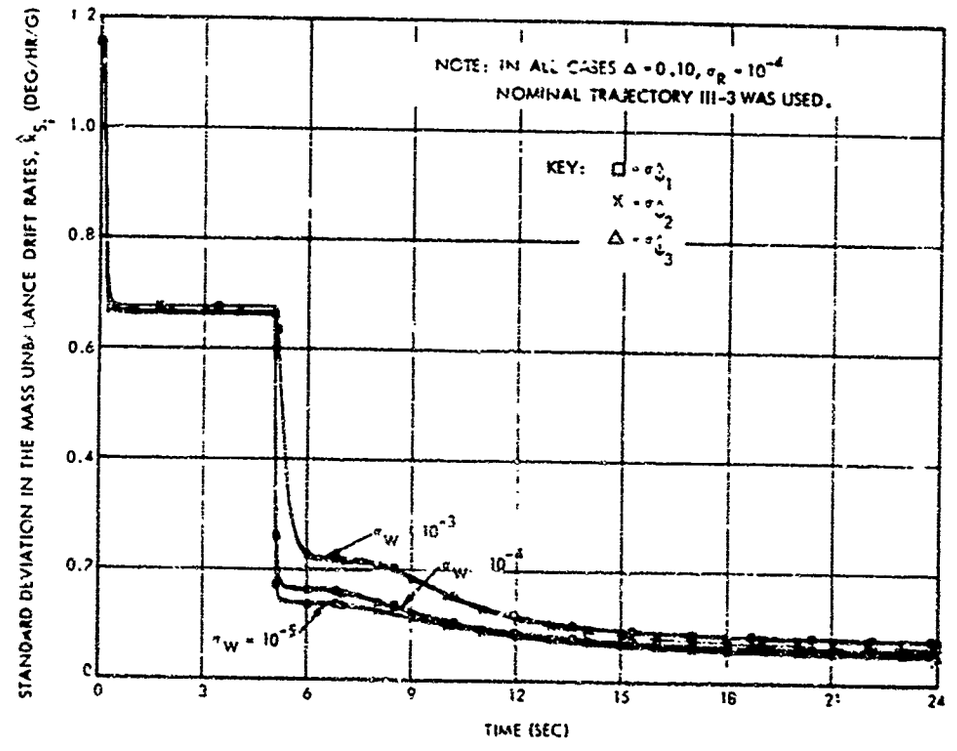


Figure 4.18 Effect of Measurement Noise on the Estimation of the Initial Misalignment Angles, ψ_i

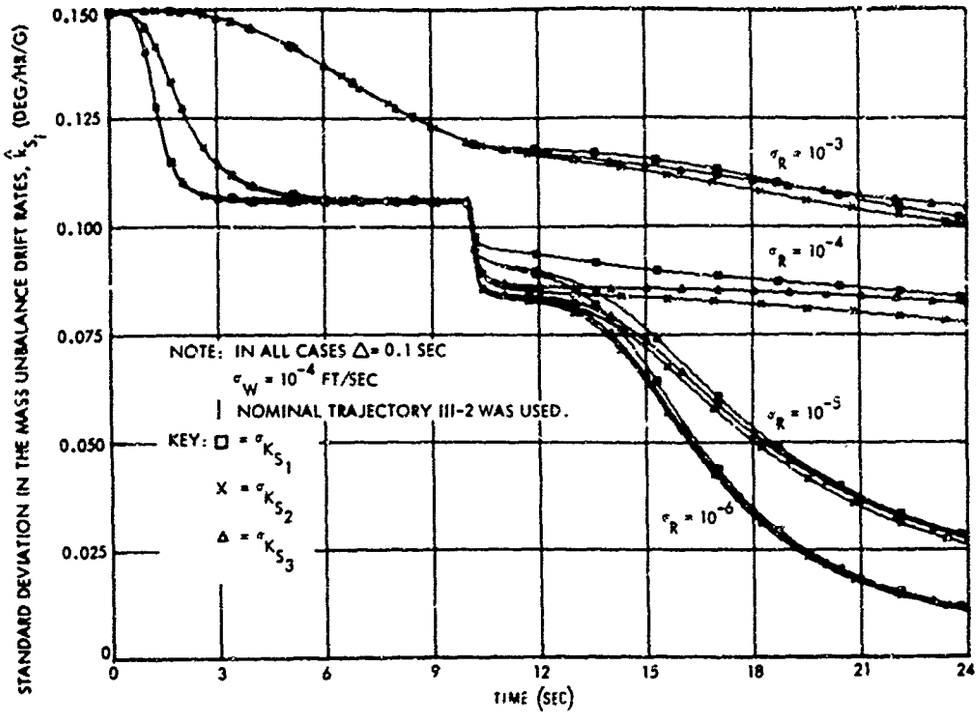


Figure 4.19a Effect of σ_R on the Estimation k_{S_i}

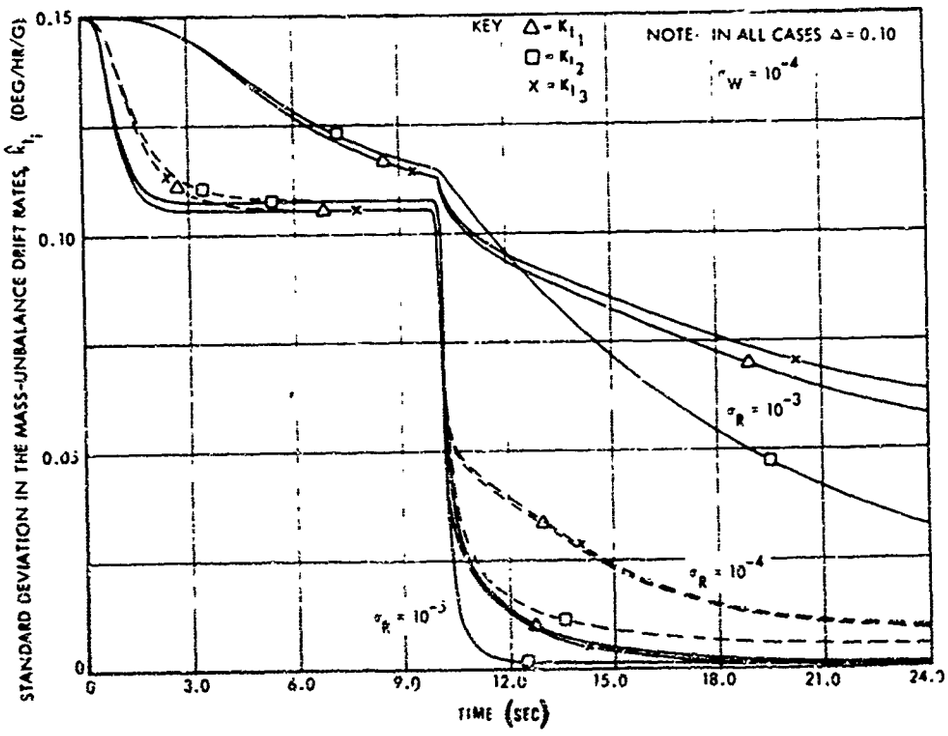


Figure 4.19b Effect of σ_R on the Estimation of k_{I_i}

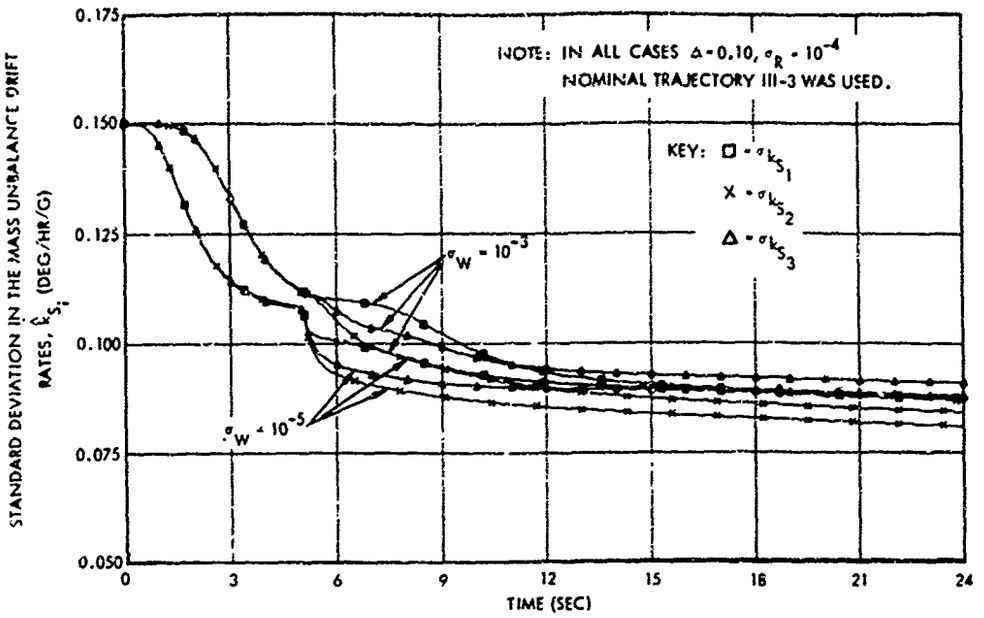


Figure 4.20a Effect of σ_W on the Estimation of k_{S_i}

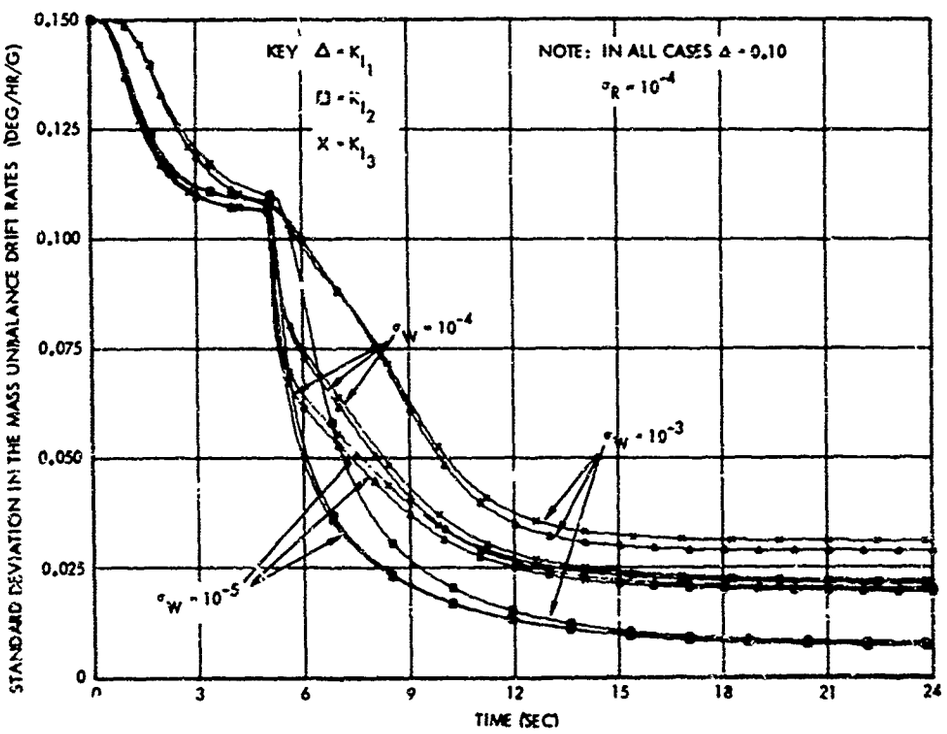
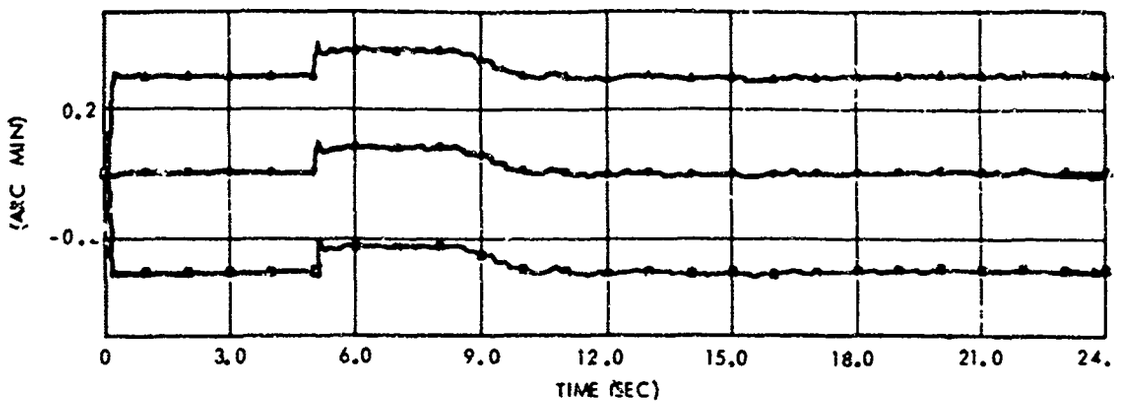
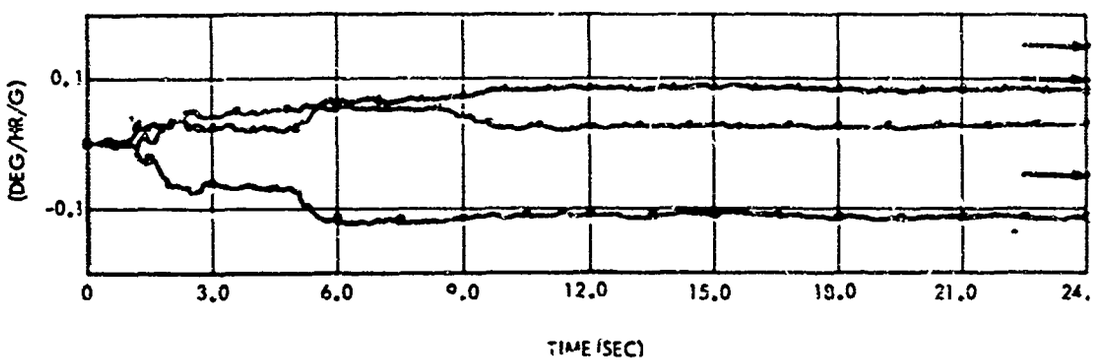


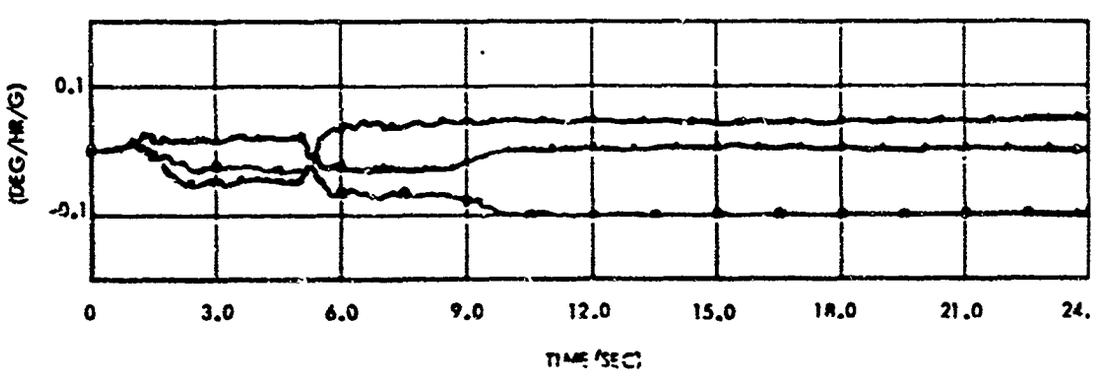
Figure 4.20b Effect of σ_W on the Estimation of k_{I_i}



a. ESTIMATES OF INITIAL MISALIGNMENT ANGLE, $\hat{\psi}_i$



b. ESTIMATE OF SPIN AXIS MASS-UNBALANCE DRIFT RATE, $\hat{k}_{S_i}$



c. ESTIMATES OF INPUT AXIS MASS-UNBALANCE DRIFT RATE, $\hat{k}_{I_i}$

Note: Nominal III-3 used (Figure 4.16) with $\sigma_R = 10^{-4}$ and $\sigma_W = 10^{-5}$. Actual error parameter values are shown by arrows.

Figure 4.21 Minimum Variance Estimates of Ψ_i, k_{S_i}, and k_{I_i}, i=1,2,3.

4.5 Discussion of the Steady State Estimates

The knowledge of the steady state estimates and of the error bounds on the estimation covariances is not required directly in the maneuver optimization problem. However, this type of discussion could help to answer certain problems of controllability which might arize. For example, the final error covariances which need to be specified for certain trajectory optimization problems (Section 5.4) could not be less than the steady state results given below. For this reason some results which might be applied to the asymptotic estimator of the covariances are included [4]. To present the main results of this section, the general state and measurement equations are written

$$\underline{x}_k = \Phi_{k,k-1}\,\underline{x}_{k,k-1} + \underline{w}_{k-1}$$

$$\underline{z}_k = M\,\underline{x}_k + \underline{v}_k$$

$$\text{Cov}(\underline{w}) \equiv Q \text{ and } \text{Cov}(\underline{v}) \equiv R$$

We write M instead of M_k in the last equation because $M = (I \vdots O)$, a constant matrix in the present problem.

Definition: The above system is said to be q-stage observable $(1 \le q \le N)$ on an interval $t_o \le t_k \le t_N$ if and only if the matrix $M_{k,k-q+1}$ (which is defined below) is positive definite for arbitrary t_k and q such that $t_1 \le t_{k-q+1}$ and $t_k \le t_N$.

$$M_{k,j} = \sum_{i=j}^{k} \phi_{i,k}^{*} \, M^{*} R_{j}^{-1} M \phi_{i,k}$$

In the present problem

$$\phi^{*} M^{*} R^{-1} M \phi = \sigma_{w}^{-2} \begin{bmatrix} \phi 1^{*} \phi 1 & \phi 1^{*} \phi 2 \\ \phi 2^{*} \phi 1 & \phi 2^{*} \phi 2 \end{bmatrix}$$

which can be made positive definite for the three configurations considered here.

Result 1: If the process noise is zero and if the system is q-stage observable, then the error covariance matrix of the estimates P_k vanishes as $k \to \infty$ if $\|(M_{k,1}^{M})^{-1}\| \to 0$ more rapidly than $\|\phi_{k,0}\|^{2}$ increases, where

$$M_{k,1}^{M} = \sum_{i=1}^{k} \phi_{i,0}^{*} M^{*} R_{i}^{-1} M \phi_{i,0}$$

Result 2: If the process noise is zero and the system is q-stage observable, then P_k becomes essentially independent (that is, for k large enough) of the a priori statistics P_0.

Result 3: If the measurement noise is zero, then P_k satisfies

(a) $MP_k = 0$ and $P_k M^{*} = 0$ for each k,

(b) P_k is nonnegative definite, but never positive definite

(c) If $P_k' \equiv \phi_{k,k-1}^{*} P_{k-1} \phi_{k,k-1} + Q_{k-1}$ is positive definite, and $m<n$ (where M is $m \times n$), then P_k can never vanish.

Further, the error in the estimate $\tilde{\underline{x}}_k = \hat{\underline{x}}_k - \underline{x}_k$ is any element of the null space of M, i.e., $M\tilde{\underline{x}}_k = \underline{0}$ for arbitrary k.

Result 4: The error bounds for P_k for the q-stage observable system are given by

$$\Phi_{k,0}\left[P_0^{-1} + M_{k,1}^{M}\right]^{-1}\Phi_{k,0}^{*} + P_k^{P}$$

$$\leq P_k \leq \Phi_{k,0}\left[P_0^{-1} + M_{k,i}^{N}\right]^{-1}\Phi_{k,0}^{*} + W_{k,1}^{S}$$

where P_k^P is any nonnegative-definite matrix (never positive definite and generally nonzero) and

$$W_{k,1}^{S} \equiv \sum_{i=1}^{k} \Phi_{k,i}\, Q_{i-1}\, \Phi_{k,i}^{*}$$

It would be useful to compute the above error bounds as a function of the various system parameters.

If it is assumed that estimator IB is q-stage observable, then the above results may be used to interpret Figures 4.4 and 4.5. Figure 4.4 illustrates that all of the curves seem to be going to the same steady-state values for σ_R fixed and varying σ_W. This observation is implied by Result 1 above. Result 2 implies that the estimation covariance will become essentially independent of the initial statistic P although no experimental results are presented as verification.

Figure 4.5 indicates that the $\sigma_{\hat{\Psi}_i}$ will reach steady-state values which are proportional to σ_R. This is implied from Result 3(c) above. Figure 4.4 indicates that the estimates are not improved for decreasing σ_W if the estimation time is sufficiently long. These results are not used to discuss Figures 4.9, 4.10 and Figures 4.17, 4.18, 4.19, 4.20, because steady-state values were not reached for these configurations. It should be noted that the bounds on the covariance equation which are given in [4] for the discrete case are similar to those given in Section 15 and 16 of [5] for the continuous case.

Remark 4.1 The experimental results of this chapter indicate that significant improvements in the identification of the specific error parameters considered can be obtained by changing the nominal trajectory. Furthermore, certain changes result in estimation variances which decrease quicker than for other changes. The best time to make a trajectory change for a particular error parameter seems to be when the σ for this error is approaching a steady state value. Section 4.5 indicates that if the system satisfies certain properties the steady values (that is, $t \longrightarrow \infty$) can be estimated independently of the nominal trajectory. In section 5.1 it will be seen that some of these properties are related to the "no-noise" situation.

CHAPTER V

ANALYSIS OF TRAJECTORY MANEUVER AND FORMULATION OF THE OPTIMAL CONTROL PROBLEM

5.1 Identification of errors under ideal conditions

For the purposes of determining the minimum number of trajectory changes for the uniqueness of the solutions, we consider the minimum number of trajectory maneuvers which are required to identify the IMU error parameters under ideal conditions, that is, under the conditions of no random disturbances ($\sigma_R = \sigma_W = 0$). To make the discussion as uncomplicated as possible, the simplest configuration (Configuration IB) is considered first. For the no-noise situation we have

$$\left[v_s^{(s)}(k+1,k)\right]\underline{\Psi} = k_\Psi^{-1}\left(\underline{v}^{(o,s)}(k+i) - \underline{v}^{(o,s)}(k)\right) \quad (5.1a)$$

If all subscripts are dropped and k takes on two different values, the above equation gives

$$A_1\underline{\Psi} = \underline{v}_1 \quad \text{and} \quad A_2\underline{\Psi} = \underline{v}_2 \quad (5.1b)$$

where the subscripts 1 and 2 correspond to k_1+1 and k_2+1 respectively.

Property 5.1 For the no-noise situation, at least one trajectory maneuver must be made in order that the misalignment angles Ψ_i be determined uniquely.

Demonstration The matrix A_i is of the skew symmetric form

$$A_i = \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

for which the eigenvalues are $\lambda_1 = 0$, $\lambda_2 = +j\|\underline{a}\|$ and $\lambda_3 = -j\|\underline{a}\|$, where $\|\underline{a}\|$ is the magnitude of the sensed acceleration. This means the three dimensional space in which $\underline{\Psi}$ is defined, E_3, can be spanned by the set $\{\underline{x}_i^1, \underline{x}_i^2, \underline{x}_i^3\}$, where $\underline{x}_i^j$ is the eigenvector corresponding to the eigenvalue λ_j^i, and any vector $\underline{x}$ in E_3 can be represented by

$$\underline{x} = \alpha_1 \underline{x}^1 + \alpha_2 \underline{x}^2 + \alpha_3 \underline{x}^3 = \alpha_1 \underline{x}^1 + \alpha_2 \underline{x}^2 + \alpha_3 \underline{\bar{x}}^2$$

so that

$$A\underline{\Psi} = \lambda_1 \alpha_1 \underline{x}^1 + \lambda_2 \alpha_2 \underline{x}^2 + \bar{\lambda}_2 \alpha_3 \underline{\bar{x}}^2 = \lambda_2 \alpha_2 \underline{x}^2 + \bar{\lambda}_2 \alpha_3 \underline{\bar{x}}^2 \tag{5.1d}$$

If $\underline{\Psi}$ has a component in the null space of $\lambda_1 = 0$, N_0 (i.e., $N_0 \equiv \{\underline{x} \text{ in } E_3 \mid A\underline{x} = \underline{0}\}$), then $\underline{\Psi}$ cannot be solved for. On the other hand, if $\underline{\Psi}$ does not have a component in N_0, then any multiple of a vector in N_0 can be added to $\underline{\Psi}$ and this will also satisfy the

equation $A\underline{\Psi} = \underline{v}$.

For convenience, the common factor which takes care of the thrust acceleration can be factored from the A matrix, so that we have

$$A_i \underline{\Psi} = \underline{y}_i \tag{5.2a}$$

where

$$A_i = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \text{ and } \sum_{j=i}^{3} t_j^2 = 1 \tag{5.2b}$$

The eigenvalues of A_1 and A_2 are then the same, namely,

$$\lambda_1 = 0, \quad \lambda_2 = \sqrt{-1} \text{ and } \lambda_2 = \sqrt{-1}$$

If the estimation were performed at the same time, then the problem is the following: find the vector $\underline{\Psi}$ such that

$$\left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \underline{\Psi} = \left[\begin{array}{c} \underline{y}_1 \\ \hline \underline{y}_2 \end{array}\right] \tag{5.2c}$$

If the matrix $\left[A_1 \,\vdots\, A_2\right]^*$ has rank 3, then a unique $\underline{\Psi}$ may be obtained by multiplying Equation (5.2c) and then computing $\underline{\Psi}$ by the equation

$$\underline{\Psi} = \left(\left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]^* \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \right)^{-1} \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]^* \left(\begin{array}{c} \underline{y}_1 \\ \hline \underline{y}_2 \end{array} \right) \tag{5.2d}$$

When the estimation is performed recursively, $A_1\underline{\Psi} = \underline{y}_1$ must be solved for $\underline{\Psi}$ first, and then $A_2\underline{\Psi} = \underline{y}_2$ is solved. The vector $\underline{\Psi}$ may be represented by the eigenvectors

$$\underline{\Psi} = \alpha_1^1 \underline{x}_1^1 + \alpha_2^1 \underline{x}_1^2 + \alpha_3^1 \overline{\underline{x}_1^2} = \alpha_1^1 \underline{x}_1^1 + \alpha_2^1 \underline{x}_1^2 + \overline{\alpha_2^1 \underline{x}_1^2} \quad (5.3a)$$

The vector $\underline{y}_1$ is then given by

$$\underline{y}_1 = \sqrt{-1}\left(\alpha_2^1 \underline{x}_1^2 - \overline{\alpha_2^1 \underline{x}_1^2}\right) = -2\,\mathrm{Im}\left(\alpha_2^1 \underline{x}_1^2\right) \quad (5.3b)$$

The coefficients α_2^1 may be solved for in terms of $\underline{y}_1$ and $\underline{x}_1^2$, which are both known. This corresponds to the solution of minimum norm (pseudo-inverse). If this solution is denoted by $\underline{\Psi}_1$, then $\underline{\Psi} = \underline{\Psi}_1 + c_1\underline{x}_1^1$ where c_1 is an unknown constant and $\underline{x}_1^1$ is the eigenvector of A_1 corresponding to the zero eigenvalue. Next the Equation $A_2\underline{\Psi} = \underline{y}_2$ is solved for a corresponding $\underline{\Psi}_2$, and $\underline{\Psi}$ is then given by $\underline{\Psi} = \underline{\Psi}_2 + c_2\underline{x}_1^2$. These Equations can be solved for c_1 and c_2, and thus give $\underline{\Psi}$ uniquely.

Remark 5.1 When random distrubances are added, a similar discussion goes through with the reasoning that the change in μ(and therefore the change in the matrix A) must be large enough so that the error due to $\underline{\Psi}$ can be detected over the errors due to the random disturbances.

Remark 5.2 The problem of uniqueness of the estimates for Configuration II can be considered as in the above discussion. In this case the matrices A_i are 3×6 instead of 3×3, and it is required that $[A_1^* \mid \cdots \mid A_n^*]^*$ have rank 6. The matrix A_i is the 3×6 matrix corresponding to the i^{th} observation, and is of the form

$$\left[\begin{array}{c:c} \alpha \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} & \beta \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \end{array}\right]$$

where α, β, and the t_i change with time. With no trajectory maneuvers, the matrix A^*A, where $A^* = [A_1^* \mid A_2^*]$ has two eigenvalues which are essentially zero and one which is close to zero, even though the factors α, β change at different rates. With one maneuver, the matrix A^*A still has the same zero eigenvalue properties, and with two trajectory maneuvers, the matrix A^*A, where $A^* = [A_1^* \mid A_2^* \mid A_3^*]$, has no zero eigenvalues. Thus it is sufficient for configuration II to have two trajectory maneuvers in order that $\underline{\Psi}$ and $\underline{\varepsilon}$ be identifiable under ideal conditions.

Remark 5.3 The uniqueness property for Configuration III is again discussed as above, where the matrix A_i is now of the form

$$\left[\begin{array}{c:cc} \alpha\begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} & \beta\begin{bmatrix} 0 & -t_{32} & t_{23} \\ t_{31} & 0 & -t_{13} \\ -t_{21} & t_{12} & 0 \end{bmatrix} & \begin{bmatrix} 0 & -t_{33} & t_{21} \\ t_{32} & 0 & -t_{11} \\ -t_{22} & t_{13} & 0 \end{bmatrix} \end{array}\right]$$

where α, β, t_i, t_{ij}, change with time. In this case it is required that $[A_1^* \,|\, \cdots \,|\, A_n^*]^*$ be of rank 9. For this configuration A^*A has zero eigenvalues when no maneuver is executed, and no zero eigenvalues when three maneuvers are executed.

From Property 5.1 and the above remarks, the following conjecture is made

Conjecture 5.1 In order that the parameters of a particular parameter identification problem be identifiable under ideal conditions it is sufficient that a minimum number of trajectory maneuvers be executed. The number of trajectory variations can be obtained from the transition matrix of the system. In the presence of random disturbances it is plausible that a similar situation prevails, with the maneuvers being large enough to detect the changes in the errors due to the error parameters over that due to the random disturbances.

Remark 5.4 The lower bounds of the minimum number of trajectory maneuvers could be useful in determining the "optimal

controls" discussed below and in Chapters VI and VII. In Chapter VI it is shown that the optimal controls are often bang-bang. A lower limit on the number of switches in the controls reduces the search procedure which might be used to obtain good initial guesses for the computation of the optimal controls as discussed in Chapter VII.

Figures 5.1 and 5.2 illustrate the standard deviations in the estimated misalignment angles and the corresponding estimates, Ψ_i, respectively, for $\sigma_R = \sigma_W = 0$. The trajectory pitch angle profiles, μ, for these two cases were such that μ = constant (IB-1) and μ changed discretely downward by 5 degrees at 0.2 seconds (IB-2). The "squares" in Figure 5.1b indicate the estimates for the second trajectory, where the actual values for Ψ are (0.5, 0.1, -0.3). The variances and the estimates for Configuration II are shown in Figures 5.2a through 5.2b. With no process and measurement noise, the standard deviations in the estimated misalignments, $\sigma_{\hat{\Psi}_i}$, do not drop immediately to zero, but seem to steady out at approximately 0.67. At t = 0.4 sec the pitch profile changes discretely by 5 deg, and these sigmas then go essentially to zero. The estimated misalignment angles, $\hat{\Psi}_i$, are estimated perfectly after t = 0.4 sec. The reason that these angles are not given correctly before t = 0.4 sec is that the estimated misalignments need

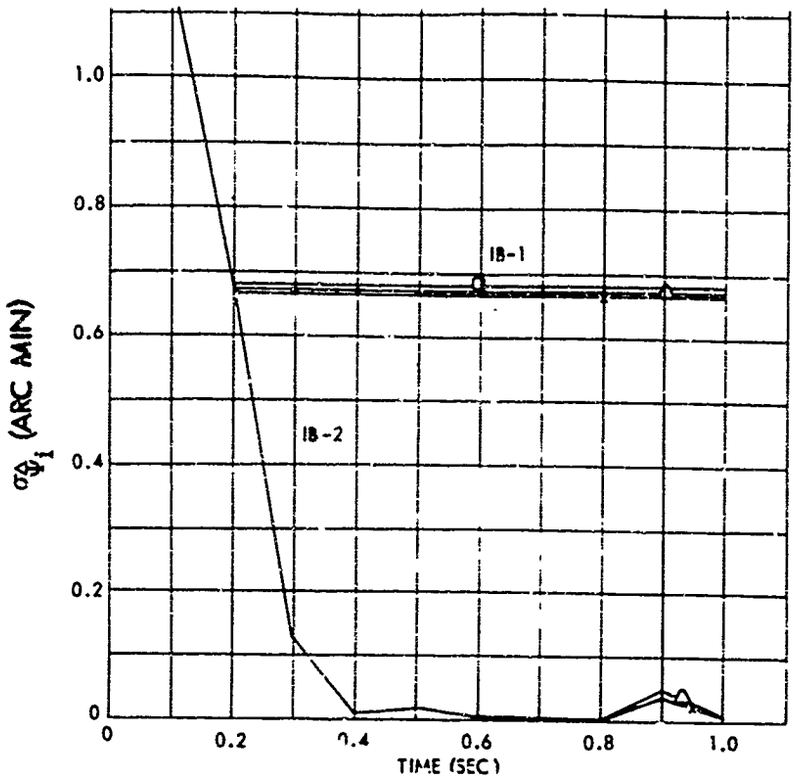


Figure 5.1a Initial Misalignment Sigmas, $\sigma_{\hat{\Psi}_i}$, for the No-Noise Situation $\sigma_R = \sigma_W = 0$

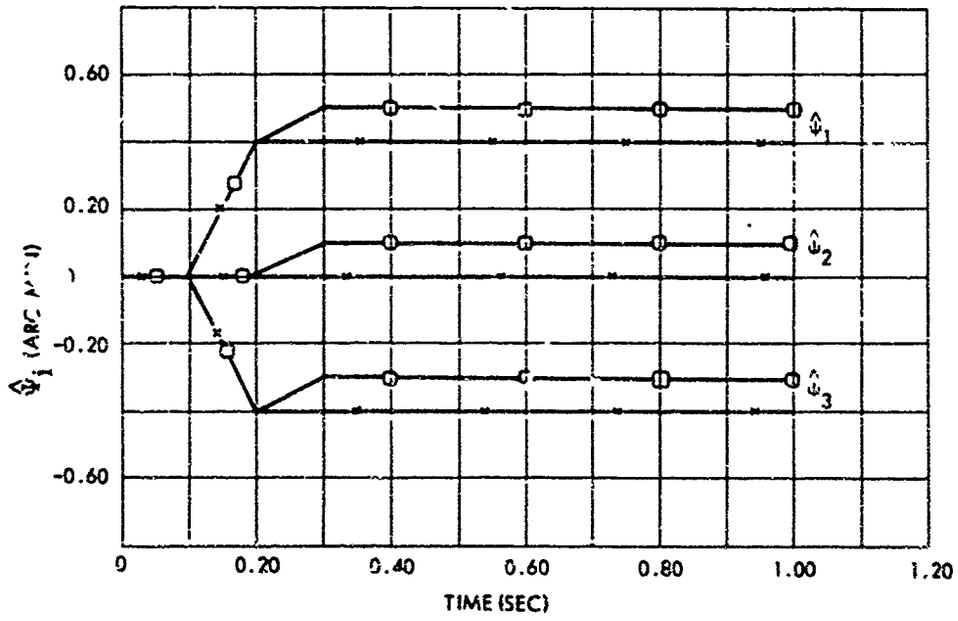


Figure 5.1b Estimates of the Misalignment Angles, $\hat{\Psi}_i$, for the No-Noise Situation $\sigma_R = \sigma_W = 0$ (actual $\underline{\Psi}$ values are 0.5, 0.1, and -0.3).

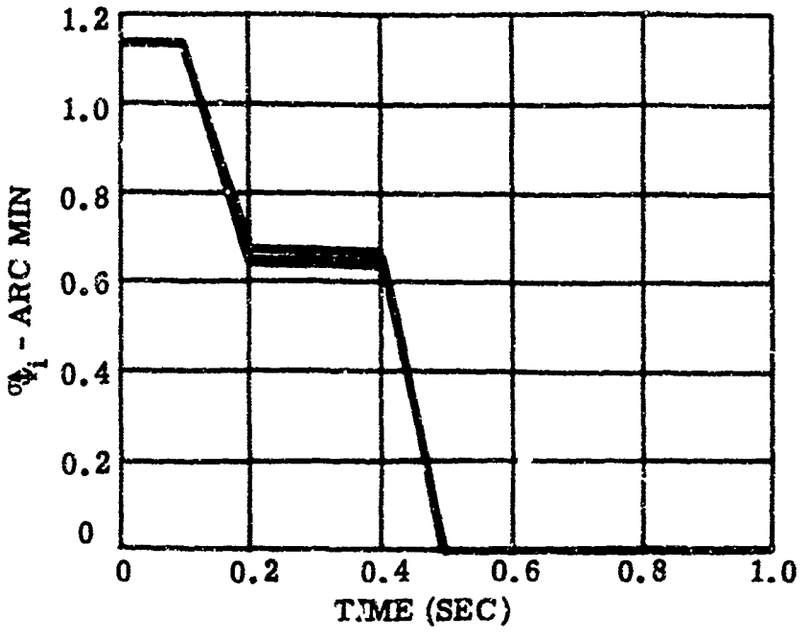


Figure 5.2a Misalignment Sigmas, $\sigma_{\hat{\Psi}_i}$, for $\sigma_R = \sigma_W = 0$.

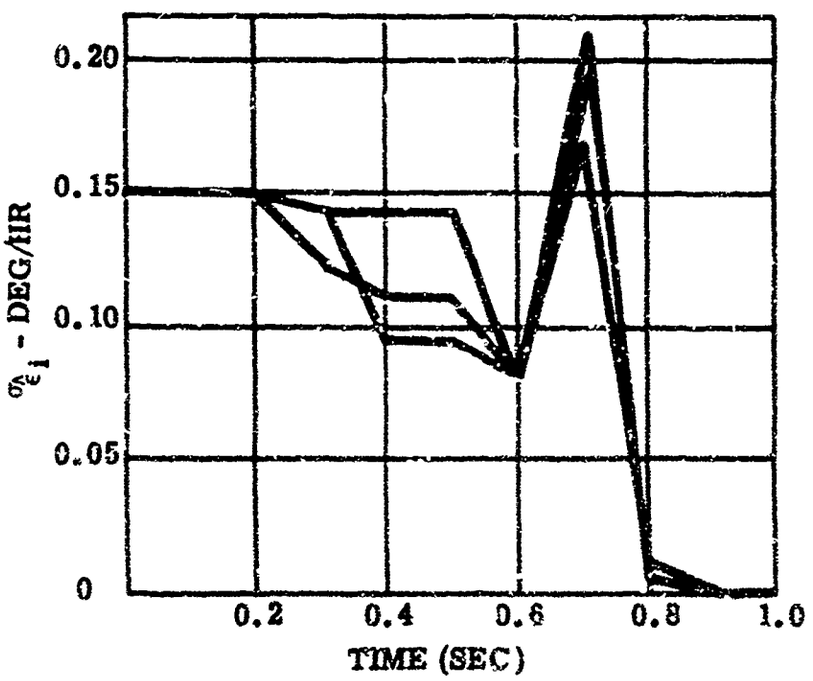


Figure 5.2b Constant Gyro Drift Sigmas, $\sigma_{\hat{\epsilon}_i}$, for $\sigma_R = \sigma_W = 0$.

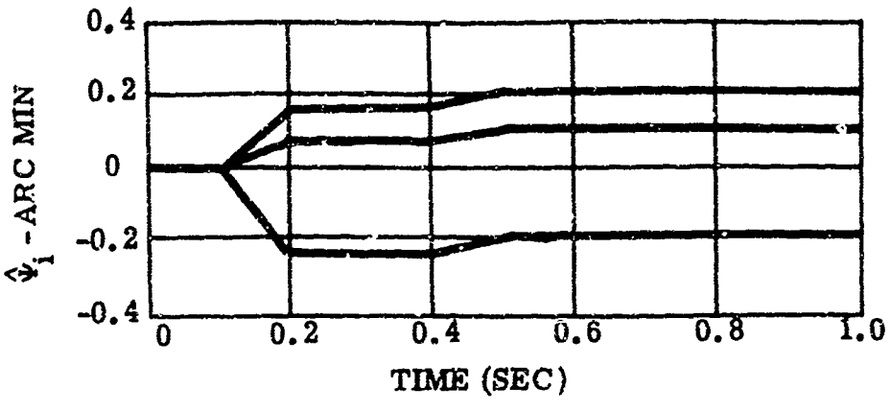


Figure 5.2c Estimates $\hat{\Psi}_i$ for the case $\sigma_W = \sigma_R = 0$
(Actual Values were 0.1, 0.2 and -0.2)

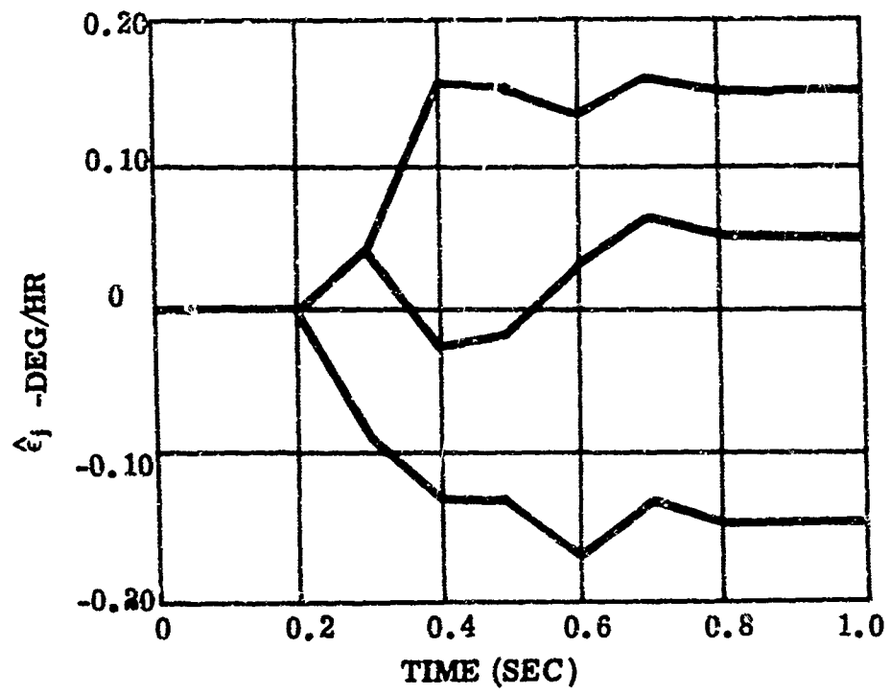


Figure 5.2d Estimates $\hat{\epsilon}_i$ for the case $\sigma_R = \sigma_W = 0$
(Actual Values were 0.15, 0.05 and -0.15)

not be unique. With the estimation of the constant drift rates, ϵ_i, there is some difficulty at first. At 0.7 sec it seems that there is a "numerical instability" in the $\sigma\hat{\epsilon}_i$. However, the sigmas settle down to essentially zero at t = 1.0, and the estimated ϵ_i come out perfectly.

Using the nominal trajectories shown in Figure 5.3 below (Nominals III-1a and III-1b), the sigmas for the misalignment angles, $\sigma\hat{\Psi}_i$, and the sigmas for the mass-unbalance gyro drift rates, $\sigma\hat{k}_{S_i}$ and $\sigma\hat{k}_{I_i}$, for the no-noise conditions ($\sigma_R = \sigma_W = 0$) are shown in Figures 5.4 and 5.5 respectively. The corresponding estimates $\hat{\Psi}_i$, $\hat{k}_{S_i}$, and $\hat{k}_{I_i}$, are shown in Figure 5.6. It is noted in Figures 5.4 and 5.5 that the sigmas do not go to zero for the constant pitch profile (TP III-1a). When the pitch profile is varied (TP III-1b), the sigmas go down to zero. For this particular case perfect estimates of Ψ_i are obtained for both nominal trajectories (Figure 5.6).

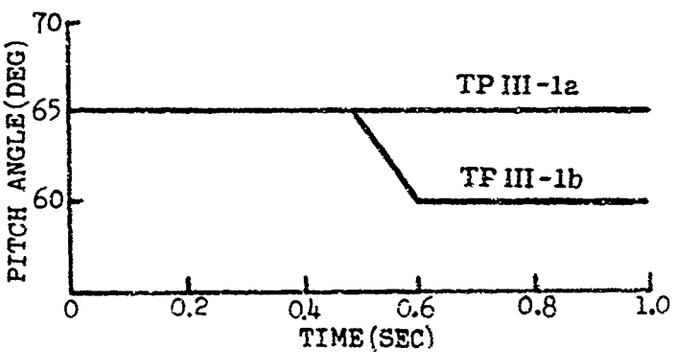


Figure 5.3 Nominal Pitch Profiles III-1a and III-1b

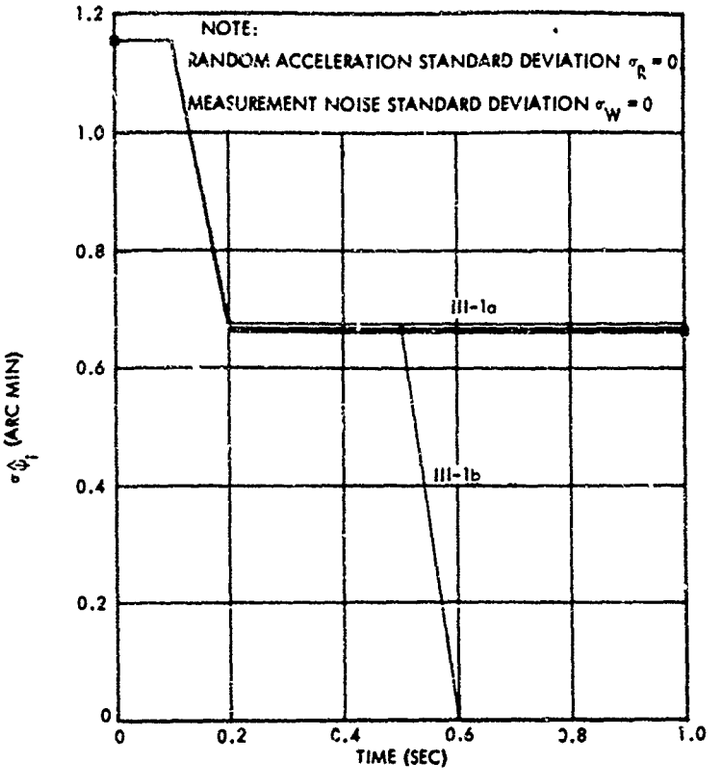


Figure 5.4 Misalignment Sigmas, $\sigma_{\hat{\psi}_i}$, Using Nominals III-1a and III-1b with $\sigma_R = \sigma_W = 0$.

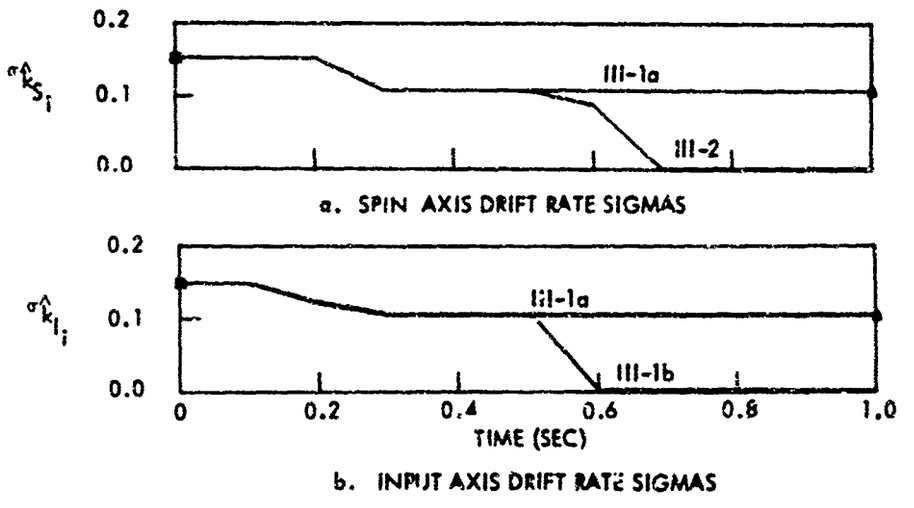
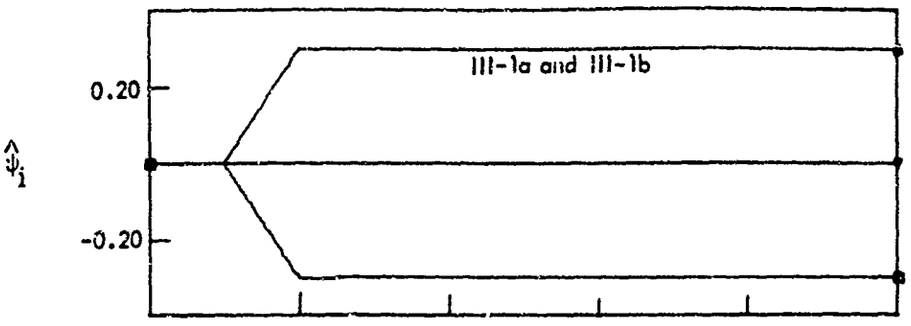
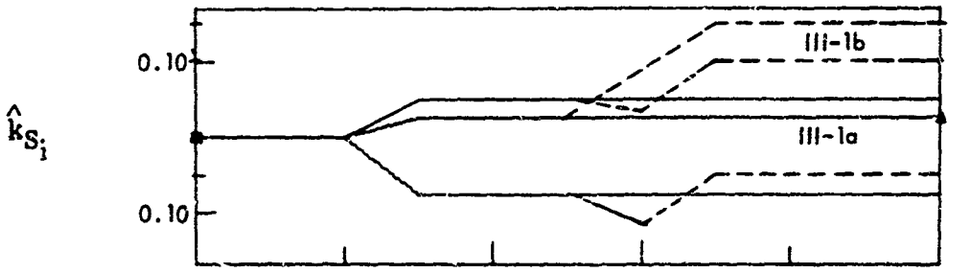


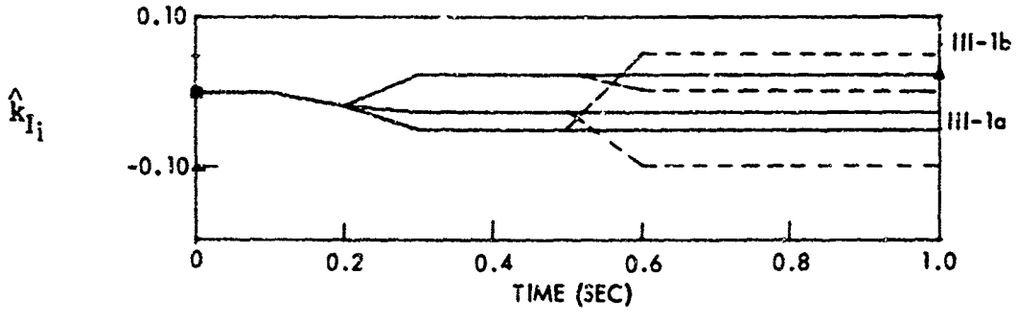
Figure 5.5 Mass-Unbalance Drift Rate Sigmas, $\sigma_{\hat{k}_{Si}}$ and $\sigma_{\hat{k}_{Ii}}$ using Nominals III-1a and III-1b with $\sigma_R = \sigma_W = 0$.



a. MISALIGNMENT ANGLE ESTIMATES, $\hat{\psi}_i$ (ACTUAL VALUES 0.3, 0, -0.3)



b. SPIN AXIS DRIFT RATE ESTIMATES, $\hat{k}_{S_i}$ (ACTUAL VALUES 0.15, 0.10, and -0.05)



c. INPUT AXIS DRIFT RATE ESTIMATES, $\hat{k}_{I_i}$ (ACTUAL VALUES 0.05, 0, -0.10)

Figure 5.6 **Parameter Estimates $\hat{\psi}_i$, $\hat{k}_{S_i}$ and $\hat{k}_{I_i}$ Using Nominals III-1a and III-1b with $\sigma_R = \sigma_W = 0$**

5.2 Formulation of the Optimal Control Problem

In this section the manner in which the control (trajectory pitch angle) enters the covariance equation for the error parameters $\underline{\alpha}$ is indicated, and the pertinent optimization equations are formulated. Under the assumptions for Configurations IB, II and III, the state and observation equations are, respectively

$$\dot{x} = Ax + Bw$$

and

$$z = Mx + v$$

where $x = (v^{(o,s)}, \alpha)^*$, a $(3+N) \times 1$ vector, where N indicates the number of components in α, $\alpha_I = \underline{\Psi}$ (Configuration I), $\alpha_{II} = (\underline{\Psi}, \underline{\epsilon})^*$ (Configuration II), $\alpha_{III} = (\underline{\Psi}, \underline{k}_S, \underline{k}_I)^*$ (Configuration III), w is a 3×1 vector of random (uncorrelated) vehicle accelerations, and v is a 3×1 vector of uncorrelated observation disturbances. The respective covariances of these random vectors are

$$Q = \text{cov}(w) = q \cdot I \quad \text{and} \quad R = \text{cov}(v) = r \cdot I$$

The more convenient notation q and r are used from here on instead of σ_R^2 and σ_W^2. The matrix B is defined by $B = \begin{pmatrix} T^{(o,s)} \cdot \underline{T}^{(m,o)} \\ O \end{pmatrix}$ $\left((3+N) \times 3\right)$, where $T^{(o,s)}$ is the 3×3 coordinate transformation from the master to the slaved IMU. This transformation involves the angles which orient the slaved IMU, $\theta_i^{(o,s)}$, $i = 1, 2, 3$, which

are now assumed to be fixed for the "equal acceleration" orientation and $T^{(m,o)}$ is the 3x3 coordinate transformation from the vehicle to the master IMU coordinates. This transformation involves the trajectory pitch profile, $\theta_2^{(o,m)}$, which we write as

$$\theta_2^{(o,m)}(t) \equiv \mu_o(t) + \mu(t)$$

to denote a nominal profile $\mu_o(t)$ and an off-nominal control in the profile $\mu(t)$. The observation matrix M is given by

$$M = \left[\begin{array}{c:c} I & O \end{array}\right] \quad (3 \times (3+N))$$

The matrix A is given by

$$A = \left[\begin{array}{c:c} O & A_i \\ \hdashline O & O \end{array}\right], \quad i = I, II, III,$$

where

$$A_I = k_\Psi \left[a_s^{(s)}\right] \qquad (3 \times 3)$$

$$A_{II} = \left[k_\Psi \left[a_s^{(s)}\right] \,\vdots\, k_{d.r.}\, t\left[a_s^{(s)}\right]\right] \qquad (3 \times 6)$$

$$A_{III} = \left[k_\Psi \left[a_s^{(s)}\right] \,\vdots\, k_{m.u.} \left[a_s^{(s)}\right] \int_0^t \left[a_s^{(s)}\right]_2 d\tau\right] \qquad (3 \times 9)$$

It is recalled that the matrix $\left[a_s^{(s)}\right]$ is defined by

$$\left[a_s^{(s)}\right] \equiv \begin{pmatrix} 0 & -a_{s_3}^{(s)} & a_{s_2}^{(s)} \\ a_{s_3}^{(s)} & 0 & -a_{s_1}^{(s)} \\ -a_{s_2}^{(s)} & a_{s_1}^{(s)} & 0 \end{pmatrix}$$

where $a_{s_i}^{(s)}$, $i = 1, 2, 3$ are the components of the nominal acceleration in the IMU_S coordinates, $\underline{a}_s^{(s)}$. The superscript and subscript s on the acceleration will be dropped, unless the meaning is not clear. In particular, each component is given by

$$a_i(t) = T_i(t)\, a_s^{(m)}(t)$$

where

$$\underline{T} \equiv \begin{bmatrix} T_1 \\ T_2 \\ T_3 \end{bmatrix} = T^{(o,s)}\ T^{(m,o)} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

For the equal acceleration orientation and $\theta_2^{(o,m)} = -25^o$, $\underline{T} \equiv \underline{T}_o = \frac{1}{\sqrt{3}}(1, 1, 1)^*$. The thrust acceleration $a_s^{(m)}(t)$ is of the form $\alpha/(\beta - t)$, which approximates a vehicle with a constant propellant burning rate.

The matrix $\left[a_s^{(s)}\right]_2$ is defined by

$$\left[a_s^{(s)}\right]_2 \equiv \left[\begin{array}{ccc:ccc} a_1 & 0 & 0 & a_2 & 0 & 0 \\ 0 & a_2 & 0 & 0 & a_3 & 0 \\ 0 & 0 & a_3 & 0 & 0 & a_1 \end{array}\right]$$

when multiplied out,

$$\left[a_s^{(s)}\right]\int_0^t \left[a_s^{(s)}\right]_2 dt \equiv \left[\begin{array}{ccc:ccc} 0 & -a_{32} & a_{23} & 0 & -a_{33} & a_{21} \\ a_{31} & 0 & -a_{13} & a_{32} & 0 & -a_{11} \\ -a_{21} & a_{12} & 0 & -a_{22} & a_{13} & 0 \end{array}\right]$$

where

$$a_{ij}(t) \equiv T_i(t)\, a_s^{(m)}(t)\int_0^t T_j(\tau)\, a_s^{(m)}(\tau)\, d\tau, \quad i,j = 1,2,3.$$

If we make a small angle approximation for $\mu(t)$, we obtain

$$\underline{a}_s^{(s)}(t) = \frac{\alpha}{\beta - t}\left\{\begin{pmatrix}\cos\mu_0 \\ 0 \\ -\sin\mu_0\end{pmatrix} + \mu(t)\begin{pmatrix}-\sin\mu_0 \\ 0 \\ \cos\mu_0\end{pmatrix}\right\}$$

For $\mu_0(t) = -25^0$ and the equal acceleration orientation,

$$\underline{a}_s^{(s)}(t) = \frac{\alpha}{\beta - t}\left\{\frac{1}{\sqrt{3}}\begin{pmatrix}1 \\ 1 \\ 1\end{pmatrix} + \mu(t)\begin{pmatrix}.815 \\ -.394 \\ -.409\end{pmatrix}\right\}$$

Thus the control $\mu(t)$ enters linearly into A_I, A_{II}, and in a complicated bilinear fashion in A_{III}. That is,

$$a_{ij}(t) = \Big(c_i(t) + d_i(t)\mu(t)\Big)\left(\int_0^t (c_j(\tau) + d_j(\tau)\,\mu(\tau)\, d\tau\right)$$

Remark It is noted that vehicles with variable thrust engines (for example, airplanes and ships) can be considered in this framework as well. In this case, the term $a_s^{(m)}$ becomes a control variable, and $\underline{a}_s^{(s)}$ is of the form

$$\underline{a}_s^{(s)} = u_1(t)\,\underline{c}_1(t) + \underline{c}_2(t)\,u_1(t)\,u_2(t),$$

where $\underline{c}_1(t)$, $\underline{c}_2(t)$ are known 3×1 vectors, $u_1(t)$ is the thrust acceleration, and $u_2(t)$ is the trajectory pitch angle.

The estimation covariance equation, $P = E[(x-\hat{x})(x-\hat{x})^*]$, is specified by

$$\dot{P} = AP + PA^* - PM^* R^{-1} MP + Q$$

If P is partitioned in the obvious way

$$P = \begin{bmatrix} P^{xx} & P^{x\alpha} \\ P^{\alpha x} & P^{\alpha\alpha} \end{bmatrix}$$

there results

$$\dot{P}^{xx} = P^{\alpha x^*} A_i^* + A_i P^{\alpha x} - r^{-1} P^{xx} P^{xx} + q \cdot I$$

$$\dot{P}^{\alpha x} = P^{\alpha\alpha} A_i^* - r^{-1} P^{\alpha x} P^{xx}$$

$$\dot{P}^{\alpha\alpha} = -r^{-1} P^{\alpha x} P^{\alpha x^*}$$

Remark If the column vectors of $P^{\alpha x}$ are written $p_1, \ldots, p_N$, then

$$P^{\alpha x} P^{\alpha x^*} = \begin{pmatrix} p_1^* p_1 & \cdots & p_1^* p_N \\ \vdots & & \vdots \\ p_N^* p_1 & \cdots & p_N^* p_N \end{pmatrix}$$

and

$$\text{Trace } P^{\alpha x} P^{\alpha x^*} = \sum_{i=1}^{N} p_i^* p_i = \sum_{i=1}^{N} \sum_{j=1}^{3} (p_{ij}^3)^2$$

Thus

$$\text{Trace } P^{\alpha\alpha}(t) = \sum_{i=1}^{N} p_{ii}^4(0) - \sum_{i=1}^{N} \sum_{j=1}^{3} \int_0^t (p_{ij}^3(t))^2 dt$$

so that minimizing Trace $P^{\infty}(t)$ is the same as maximizing

$$\sum_{i=1}^{N} \sum_{j=1}^{3} \int_{c}^{t} (p_{ij}^{3}(t))^2 \, dt$$

The maximization is performed with respect to changes in the trajectory pitch angle, $\mu(t)$.

5.3 Trajectory Constraints

The constraints on the vehicle's trajectory will determine constraints on the pitch angle μ. In the simplest forms, the trajectory constraints might be specified in several different ways as listed below.

$$|\mu(t)| \leq M_1 \qquad \text{(C-1)}$$

$$|\dot{\mu}(t)| \leq M_2 \qquad \text{(C-2)}$$

$$\int_0^T \mu(t)dt \leq M_3 \qquad \text{(C-3)}$$

$$\left|\int_0^T \mu(t)dt\right| \leq M_4 \qquad \text{(C-4)}$$

and various combinations of the above constraints. The constraint (C-4) is the most difficult to treat analytically, and will not be discussed further here. Constraint (C-2) can be considered by defining $\dot{\mu}$ as the control, and μ as a state variable. If constraint (C-1) is also in force, then the problem becomes one with bounded state variables. In this case the specification of the necessary

conditions which the optimal control μ must satisfy become quite complicated. However, with some of the cost functionals in the next section it can be shown (Chapter VI) that with (C-1) in force, we require $|\mu(t)| = M_1$ and with (C-2) in force we require $|\dot{\mu}(t)| = M_2$, that is, μ and $\dot{\mu}$ are bang-bang controls. Thus with (C-1) and (C-2) in force, then it is plausible that $\dot{\mu}(t) = \pm M_2$ until $\mu(t) = \pm M_1$ and then $\dot{\mu}$ is switched to $\mp M_2$. The constraint (C-3) gives a measure of the amount of trajectory maneuvering allowed, or required, in order that the vehicle's position and velocity fall within some specified region at time $t = T$.

Pointwise constraints of the form (C-1),(C-2) specify the control functions to belong to a set of U for each $t \in I$, whereas global constraints of the form (C-3),(C-4) specify the controls as subsets of certain function spaces. In general, the control functions are considered to be bounded, measurable functions and the control set $\mathcal{U}$ (considered as a subset of a function space) is as follows:

$$\mathcal{U} = \{u: u \text{ is bdd, mble, } u(t) \in U(t) \text{ for } t \in I\}$$

where

$$U = \{u(t): \text{(C-1) or (C-2) hold}\}$$

or else

$$\mathcal{U} = \{u: u \text{ is bdd, mble, (C-3) or (C-4) hold}\}$$

For analytical purposes, (C-3), with no pointwise constraints would be the easiest to handle. However, the cost function should then be such that meaningful expressions for u are obtained when the Hamiltonian is minimized. For example, if it is only required to minimize Trace $P^{\alpha\alpha}(t)$ with the constraint (C-3), then the control u is eliminated completely when the Hamiltonian is minimized.

5.4 Cost Functionals

Choosing the appropriate cost function is an important aspect of the design problem. When it is only required to obtain good estimates of the error parameters in some fixed interval of time, then a reasonable cost functional would be

$$J_1(\mu) = \text{Trace}\left[W^{\alpha} P^{\alpha\alpha}(T;\mu)\right], \qquad \text{(J-1)}$$

where W^{α} is a positive definite weighting matrix. More generally, all of the members of P(T) could be considered in the minimization by considering the cost functional

$$J_2(\mu) = \text{Trace}\left[W P(T;\mu)\right], \qquad \text{(J-2)}$$

where W is a weighting matrix. On the other hand, it may only be required that the diagonal elements of $P^{\alpha\alpha}$ fall within a given region in the least possible time. This could be written as

$$J_3(\mu) = \int_0^{T^*} dt \quad \text{and} \quad \underline{P}(T^*;\mu) \in G \tag{J-3}$$

where $\underline{P}(T^*;\mu)$ denotes the elements of the covariance matrix P which are to lie within the region G. The region G is assumed to have a smooth boundary. With this cost function, the question of controllability must be considered. For the minimization to make sense, it must be verified that controls exist which can drive the states $\underline{P}$ into the region G.

It is possible that it is disadvantageous to perform too much maneuvering for the purpose of identifying the error parameters. In this case a term of the form $\int_0^T |\mu(t)|\, dt$, which is proportional to the fuel used, or else a term $\int_0^T \mu^2(t)\, dt$, which is proportional to the energy used for the maneuvering, could be added to the above cost functionals. Thus we could have

$$J_4(\mu) = \text{Trace}\left[W P(T;\mu)\right] + C_p \int_0^T |\mu(t)|^p \, dt \quad p = 1, 2 \tag{J-4}$$

$$J_5(\mu) = \int_0^T dt + C_p \int_0^T |\mu(t)|^p \, dt \quad p = 1, 2 \tag{J-5}$$

The weighting factor C_p might be chosen large enough so that pointwise constraints on μ are not required. In this case the linearization which is made for μ in section 5.2 would still be valid.

In the next chapter, a particular optimization problem will be specified by referring to the various constraints (C-i) and the cost functionals (J-i).

We can now make the general definition:

> The optimization problem is to find a control $\bar{\mu}$ such that $\bar{\mu}$ is a member of the admissible class of $\mathcal{U}$, and
>
> $$J(\bar{\mu}) \leq J(\tilde{\mu}) \text{ for all } \mu \in \mathcal{U}$$

5.5 The Simplest Examples

In this section we consider the example

$$\dot{x} = \mu x + w, \quad y = mx + v \; ; \; \text{cov}(w) = q, \quad \text{cov}(w) = r,$$

to get an idea of the minimization procedure. In this case the covariance equation satisfies the scalar equation

$$\dot{p} = 2\mu p - r^{-1}m^2p^2 + q, \quad p(o) = p_o \text{ given.}$$

Problem I $\qquad J_1(\mu) = p(T), \; |\mu(t)| \leq M_1 \; ; \; T, M_1$ given.

Minimizing the Hamiltonian gives $u = -M_1 \text{ sign } (\lambda p)$, where λ satisfies the adjoint equation

$$\dot{\lambda} = -\partial H/\partial p = (-2\mu + 2r^{-1} m^2 p)\lambda \; ; \quad \lambda(T) = 1$$

If $q > 0$, then $p(t) > 0$ and $\mu = -M_1 \operatorname{sgn} \lambda_1$. It is not difficult to see that $\lambda(t) > 0$, and therefore $\mu(t) \equiv -M_1$.

Problem II $\quad J_2(\mu) = \int_0^T dt$, $p(T) \leq \alpha$, and $|\mu(t)| \leq M_1$

Minimizing $H = \lambda \cdot p + \lambda_o \cdot 1$ gives $u(t) = -M_1 \operatorname{sgn}(\lambda p)$. At $t = T^*$, the minimum time, $p(T^*) = \alpha$, $H(T^*) = 0$, and therefore

$$\lambda(T)\ (-2M_1 \operatorname{sgn}\ \lambda(T) - r^{-1} m^2 \alpha^2 + q) + 1 = 0$$

Assuming $\lambda(T) < 0$, means that $2M_1 \alpha - 2r^{-1} m^2 \alpha^2 + q > 0$, which implies $p(T^*)$ is increasing. This contradiction implies $\lambda(T^*) \geq 0$ in $[0,T]$, and thus $\mu(t) = -M_1$, as above.

For $\mu = a$, a constant, the solution to the covariance equation can be written as

$$\frac{1}{\sqrt{q m^2 r^{-1} + a^2}}\ \operatorname{Tan h}^{-1}\ \left(\frac{2r^{-1}m^2 x - 2a}{2\sqrt{q m^2 r^{-1} + a^2}}\right)\Bigg|_{p_o}^{p} = T$$

For the first problem T is given, and is required to minimize $p(T)$. In the second problem $p(T) = \alpha$ is given, and it is necessary to minimize T. From this expression it is clear that we choose $a = -M_1$.

CHAPTER VI

Existence of Optimal Controls and Necessary Conditions for Optimal Control

In this chapter the existence of the optimal controls and the necessary conditions which the optimal controls must satisfy for the optimization problems formulated in Chapter V, and for generalizations of these problems, are considered. It is of practical importance to insure the existence of the optimal controls, since the maximum principle of Pontryagin is used to obtain the necessary conditions. Matrix notation is used to specify the Hamiltonian, the adjoint equation, and the equation which gives the optimal control in terms of the covariance matrix and the matrix of adjoint variables.

Although the equations which must be satisfied along the optimal trajectories are quite difficult to solve, and the resulting controls are open loop in nature, the results are still of practical interest. This is due to the fact that in the present application to parameter identification, it is entirely acceptable to devote considerable effort to obtaining optimal trajectory maneuvers before any experimental work or "flight testing" is performed.

6.1 Existence of Optimal Controls for Problem I

For Problem I we consider a cost function of the form

$$J_1(\underline{u}) = \text{Trace } WP(T;\underline{u}) \tag{6.1}$$

where W is a non-negative weighting matrix, T is the specified time of operation for the identification procedure, and $\underline{u}$ is a control vector. The manner in which $\underline{u}$ enters the covariance equation

$$\dot{P} = A(\underline{u})P + PA^*(\underline{u}) - PM^*R^{-1}MP + Q \tag{6.2}$$

is assumed to be such that the elements of A which involve $\underline{u}$ can only be linear combinations of the components of $u = (u^1, \cdots, u^m)^*$. Further, it is assumed that for each $t \in I \equiv [t_0, T]$, $\underline{u}(t) \in U(t)$, a compact, convex set in E_m, and it is assumed that $\underline{u}$ satisfies the constraint (6.3), where $\underline{c}$ is a constant vector.

$$\int_I \underline{u}(t)\,dt = \underline{c} \tag{6.3}$$

The control set $\mathcal{U}$ is thus defined by the set of functions

$$\mathcal{U} = \left\{ \underline{u} : \underline{u}(t) \in U(t) \text{ for } t \in I,\ \int_I \underline{u}(t) = \underline{c} \right\}$$

The existence of an optimal control $\hat{\underline{u}} \in \mathcal{U}$ is next assured by the

Property 6.1 For the non linear system (6.2), with pointwise constraints $\underline{u}(t) \in U(t)$ for $t \in I$, and U(t), a compact,

convex set in E_m, and the integral constraint (6.3), there exists an optimal control $\bar{u} \in \mathcal{U}$ such that the cost functional (6.1) achieves a global minimum.

Demonstration It is first shown that there is a uniform bound for the elements of P, considered as a vector $\underline{P}$. That is, it shall be proven that

$$\| \underline{P} \|_2 \leq \gamma < \infty \tag{6.4a}$$

where the norm $\| \cdot \|_2$ is defined by

$$\| P \|_2 = \sqrt{\sum_{i,j} p_{ij}^2} = \sqrt{\text{Trace } P P'}, \tag{6.4b}$$

and the p_{ij} are the elements of the matrix P. Indeed, since $B \equiv M^* R^{-1} M$ and Q are non-negative,

$$\dot{P} \leq A(u) P + P A^*(u) + Q \tag{6.4c}$$

Let P' be a solution of the equation

$$\dot{P}' = A P' + P' A^* + Q, \; P'(t_0) = P_0 \tag{6.4d}$$

This solution can be written as (see Property 7.1, Chapter VII)

$$P'(t) = \Phi(t, t_0) P_0 \Phi^*(t, t_0) + \int_{t_0}^{t} \Phi(t, s) Q(s) \Phi^*(t, s) ds, \tag{6.4e}$$

where $\Phi(t, s)$ is a fundamental solution of the system

$$\frac{\partial \Phi(t, s)}{\partial t} = A(t) \Phi(t, s), \; \frac{\partial \Phi(t, s)}{\partial s} = -\Phi(t, s) A(s); \; \Phi(t, t) = I \tag{6.4f}$$

By hypothesis, the elements of A will be integrable functions on the interval I, so that the solutions to (6.4f) will be unique. Further, it can be concluded that P'(t) is symmetric and

non-negative for $t \in I$, since P_0 and Q possess this property. By taking into consideration the problem from which P(t) arizes, it can be assumed that P(t) is non-negative. Thus

$$0 \leq \langle x, P(t)x \rangle \leq \langle x, P'(t)x \rangle < \infty \tag{6.4g}$$

For non-negative symmetric matrices

$$\sup_{\|x\|=1} \langle x, Px \rangle = \sup_{\|x\|=1} \|\sqrt{P}\, x\|^2 = \lambda_1 \equiv \| P \|_1,$$

where λ_1 is the greatest eigenvalue of P. However,

$$\| P \|_2 = \sqrt{\sum_{i,j} p_{ij}^2} = \sqrt{\text{Trace } PP} = \sqrt{\sum_{i=1}^{n} \lambda_i^2} \leq \sqrt{n}\, \lambda_1,$$

which implies that

$$\frac{1}{\sqrt{n}} \| P \|_2 \leq \| P \|_1 \leq \| P \|_2 .$$

Since $J_1(u)$ is obviously finitely bounded from below, we can select a sequence $\{\underline{u}_{k'}\} \in \mathcal{U}$ such that $J(\underline{u}_{k'})$ decreases monotonically to inf $J(\underline{u})$, where $\underline{u} \in \mathcal{U}$. We let $\{P_{k'}\}$ denote the solutions to the Riccati equation (6.2) which correspond to the control sequence $\{\underline{u}_{k'}\}$. For convenience, Equation (6.2) is written in the vector form

$$\dot{\underline{P}} = \underline{F}(\underline{P}, \underline{u}, t) \tag{6.5a}$$

where the definition of the vector function $\underline{F}$ is derived from

Equation (6.2). Since the set U is compact and convex, $\mathcal{U}$ will be weakly sequentially compact for the compact interval $I = [t_0, T]$. Thus a subsequence $\{\underline{u}_k\}$ can be selected such that $\underline{u}_k(t) \xrightarrow{w} \bar{\underline{u}}(t)$ for $t \in I$. The solutions to (6.5a) for each control $\underline{u}_i$ are given by

$$\underline{P}_i(t) = \underline{P}_0 + \int_{t_0}^{t} \underline{F}(\underline{P}_i(\tau), \underline{u}_i(\tau), \tau)\, d\tau \qquad (6.5b)$$

Since $|F| \leq \alpha|P| + \beta$, and $|P|$ is uniformly bounded, the sequence $\{\underline{P}_k\}$ forms a uniformly bounded and equicontinuous family of functions. The theorem of Ascoli then assures us that a subsequence $\{\underline{P}_k(t)\}$ converges uniformly to some function $\bar{\underline{P}}(t)$, where

$$\bar{\underline{P}}(t) \equiv \underline{P}_0 + \lim_{k\to\infty} \int_{t_0}^{t} \underline{F}(\underline{P}_k(\tau), \underline{u}_k(\tau), \tau)\, d\tau \qquad (6.5c)$$

Similarly, since J_1 is continuous in P_k,

$$\lim_{k\to\infty} J_1(\underline{u}_k) = \lim_{k\to\infty} \text{trace } W P_k(T\,;\, \underline{u}_k) = \text{Trace } W\bar{P}(T) \qquad (6.5d)$$

It is still required to show that $\bar{\underline{P}}$ is the response to $\bar{\underline{u}}$ and that $\bar{\underline{u}} \in \mathcal{U}$. This is accomplished by the techniques which are used in existence theory. In particular, by the assumptions on the way $\underline{u}$ enters the covariance equation, $\underline{F}$ can be written as

$$\underline{F}(\underline{P}, \underline{u}, t) \equiv \underline{G}(\underline{P}, t) + H(\underline{P}, t)\,\underline{u} \qquad (6.5e)$$

and weak limits are taken to show that $\bar{P}$ is the response to $\bar{u}$, as in page 40 [6]. In this case Lebesgue's dominated convergence theorem and a theorem on almost uniform convergence and the continuity of G, H and of the partial derivatives of G, H with respect to the p_{ij} are used to show the weak convergence of $H(\underline{P}_k(t), t)$ to $H(\hat{\underline{P}}(t), t)$ where P_k is the response to $\underline{u}_k$ and $\hat{P}$ is the response to $\bar{u}$, and $\hat{P} = \bar{P}$. The compactness and the convexity of the restraint set $U(t)$ are used to show that $\bar{\underline{u}}(t) \in U(t)$ for a.e. $t \in I$. Then $\bar{\underline{u}}(t)$ is redefined on this null set so that $\bar{\underline{u}}(t) \in U(t)$ for all $t \in I$.

This implies that $J(\bar{u}) = \inf_{\underline{u} \in \mathcal{U}} J(\underline{u})$, and $\bar{\underline{u}}$ is the required optimal control.

6.2 Existence of Time Optimal Controls

The time optimal problem is formulated as follows. For the Riccati equation (6.2), find a control $\underline{u}$, where $\underline{u}$ belongs to some admissible set $\mathcal{U}$, such that certain elements of P are less than some prescribed values in minimum time. In particular, we shall require that the diagonal elements of P, p_{jj}, be less than α_j. To be specific, $\mathcal{U}$ is defined by (6.6), as in the previous section.

$$\mathcal{U} = \left\{ \underline{u} : \underline{u}(t) \in U(t),\ \int_{t_o}^{T} \underline{u}(t)\, dt = \underline{c} \right\} \qquad (6.6)$$

Property 6.2 If U(t) is a compact, convex set, then there exists a control $\bar{\underline{u}} \in \mathcal{U}$ such that $p_{jj} \leq \alpha_j$, $j = 1, \cdots, n$. in minimum time, where the p_{jj} are the diagonal elements of the covariance matrix P which satisfies the differential equation (6.2).

Demonstration First it is assured that there exist controls in $\mathcal{U}$ which transfer the matrix P from P_o to P_T in finite time intervals $[t_o, T]$ by appropriately choosing the positive numbers α_j. This assumption can be made valid by solving (6.2) for arbitrary $\underline{u} \in \mathcal{U}$ and then observing the p_{ij}'s until such time that they reach satisfactory levels α_j. The target set for P_T, call it X(T), will be compact since the vector P was shown to be uniformly bounded in the previous demonstration. We let $\mathcal{U}'$ denote the set of all controls in $\mathcal{U}$ which transfer P_o to P_T in the time $T-t_o$, and let $\bar{T}-t_o$ be the infimum. Thus the existence of a control $\tilde{\underline{u}} \in \mathcal{U}'$ which corresponds to the time $\bar{T}-t_o$ must be demonstrated.

The proof of this result is similar to that required in Property 6.1. Let $\{\underline{u}_k'\}$ be a minimizing sequence of controls from $\mathcal{U}'$,

which are defined on intervals $[t_0, T_k']$ in such a way that the sequence $\{T_k'\}$ approaches T monotonically from above. The controls defined in this way define a sphere in $L_2[t_0, T]$. Thus a weakly convergent subsequence $\{\underline{u}_k\}$ exists such that $\underline{u}_k(t) \rightarrow \bar{\underline{u}}(t)$ weakly in $L_2([t_0, T])$. Thus we must show that $\bar{\underline{u}} \in \mathcal{U}'$, which means that $\bar{\underline{u}}(t)$ must be in $U(t)$ for $t \in [t_0, T]$, $\bar{\underline{u}}$ transfers P_0 to $X(T)$ in a time $\bar{T} - t_0$. As in the proof of the preceeding property it can be shown (page 40 [6]) that $\bar{\underline{u}} \in U(t)$, the equibounded and equicontinuous responses $P_k(t)$ to the controls $\underline{u}_k(t)$ contain a subsequence which converges uniformly to $\bar{P}(t)$, which is the response to the limiting control $\bar{\underline{u}}(t)$. It is then shown that $\bar{P}(\bar{T}) \in X(\bar{T})$, and that $\inf_{\underline{u} \in \mathcal{U}} J(\underline{u}) = \bar{T} - t_0$, so that $\bar{\underline{u}}$ is the required optimal control.

6.3 Some Basic Properties of the Riccati Equation

In order to discuss the existence and uniqueness properties of the Riccati equation (6.7a), it is convenient to write it in the form (6.7b) or (6.7c), where the definition of $F(P, u, t)$ and $F_{ij}(P, u, t)$ is clear from (6.7a).

$$\dot{P} = A(u)P + PA(u)^* - PM^*R^{-1}MP + Q, \quad P(t_0) = P_0 \qquad (6.7a)$$

$$\dot{P} = F(P, u, t) \qquad (6.7b)$$

$$\dot{P}_{ij} = F_{ij}(P, u, t), \quad i, j = 1, \cdots, n \qquad (6.7c)$$

Since the matrix A involves the controls $\underline{u}$, it can usually only be assumed that the F_{ij} are integrable in the time interval $[t_o, T]$. The basic existence and uniqueness results for the above systems may be found in [6], [7], the Appendix of [8], and [9]. The following result can be applied directly to the above problem.

Theorem Assume that the functions F_{ij} in (6.7c) and the partial derivative $\partial F_{ij}/\partial p_{k\ell}$ are continuous in P, u, t, that is, in the space E_{n^2+m+1}. Then given an initial point $P_o = P(t_o)$, where $t_o \in I \subset E_1$, and a measurable control u, with $u(t) \in U(t)$, a set in E_m, for $t \in I$, there exists a unique absolutely continuous solution of (6.7c) on some subinterval I' of I, such that $P(t) = P_o$.

If there exist integrable functions $M(t)$ and $K(t)$ on $[t_o, T]$ such that

$$|F_{ij}(P,u,t)| \leq M(t) \quad \text{and} \quad |\partial F_{ij}(P,u,t)/\partial p_{k\ell}| \leq K(t) \tag{6.7d}$$

for $i, j, k, \ell = 1, \cdots, n$, and the solution $P(t)$ with $P(t_o) = P_o$ is uniformly bounded, that is, $\|P(t)\| \leq \gamma < \infty$ for $t \in [t_o, T]$, then this is sufficient to insure that the absolutely continuous solution $P(t)$ of (6.7) is unique for the interval $[t_o, T]$.

End of Theorem

The continuity of the functions F_{ij} and $\partial F_{ij}/\partial p_{k\ell}$ in (6.7c) with respect to the variables P, u, and t is clear. They are continuous with respect to P because the right side of (6.7c) is quadratic in P, they are continuous in u because u enters the elements of A linearly. They are continuous in t because M, R, and Q are assumed continuous in t. In order to demonstrate the condition (6.7d) it is clear that it will be sufficient to show that the elements of P(t) are uniformly bounded for $t\in[t_0,T]$, that is, $|p_{ij}(t)|\leq\gamma<\infty$ for $i,j=1,\cdots,n$, and $t\in[t_0,T]$. In the case that U(t) is a compact set, a bound of the form used in the proof of Property 6.1 can be used. Sharper bounds can be obtained by considering the equation (6.7d) below [10] instead of (6.4c) where S is a symmetric matrix which can be chosen to lower the bound P'(t) of P(t). The solution of (6.7d) is given by (6.4e).

$$\dot{P}' \leq (A-SB)P'+P'(A^*-BS)+(Q+SBS) \qquad (6.7d)$$

Kalman [5] has derived upper and lower bounds for P(t) using observability and controllability properties of the linear estimation problem. These properties, which are defined next, are for the linear estimation problem, and are not to be confused with similar notions which might be introduced for the control problem.

Definitions: The system

$$\begin{aligned} \dot{y} &= Ay + w \quad ; \quad \mathrm{cov}(w) = Q \\ z &= My + v \quad ; \quad \mathrm{cov}(v) = R \end{aligned} \tag{6.8a}$$

is said to be completely observable on $[t_0, T] \equiv I$ if

$$M(t_0, T) \equiv \int_I \Phi^*(t,T) M^*(t) R^{-1}(t) M(t) \Phi(t,T) dt \tag{6.8b}$$

is positive definite on I. The system is said to be completely controllable on I if

$$W(t_0,T) \equiv \int_I \Phi(t,T) Q(t) \Phi^*(t,T) dt \tag{6.8c}$$

is positive definite on I. The system is said to be uniformly completely observable (u.c.o.) if there exists fixed positive constants σ_1, α_1, β_1, such that

$$0 < \alpha_1 I \leq M(t-\sigma_1, t) \leq \beta_1 I \qquad \forall\, t \tag{6.8d}$$

The system is uniformly completely controllable (u.c.c.) if there exists constants σ_2, α_2, β_2, such that

$$0 < \alpha_2 I \leq W(t-\sigma_2, t) \leq \beta_2 I \qquad \forall\, t \tag{6.8e}$$

For the parameter identification problem

$$\begin{aligned} \dot{x} &= A^x + C^\alpha \alpha + w^x, & \mathrm{cov}(w^x) &= Q^x \; ; \\ \dot{\alpha} &= w^\alpha, & \mathrm{cov}(w^\alpha) &= Q^\alpha \; ; \\ z &= x + v, & \mathrm{cov}(v) &= R \; ; \end{aligned} \tag{6.9a}$$

we obtain

$$\Phi = \begin{bmatrix} \phi^x & \phi^\alpha \\ 0 & I \end{bmatrix} \qquad M = [I \mid 0] \quad ;$$

$$\Phi^* M^* R^{-1} M \Phi = \begin{bmatrix} \phi^{x*} R^{-1} \phi^x & \phi^{x*} R^{-1} \phi^\alpha \\ \phi^{\alpha*} R^{-1} \phi^\alpha & \phi^{\alpha*} R^{-1} \phi^\alpha \end{bmatrix} \tag{6.9b}$$

$$\Phi Q \Phi^t = \begin{bmatrix} \phi^x Q^x \phi^{x*} + \phi^\alpha Q^\alpha \phi^{\alpha*} & \phi^\alpha Q^\alpha \\ Q^\alpha \phi^{\alpha*} & Q^\alpha \end{bmatrix} \tag{6.9c}$$

It is thus required that the integrals of the above two matrices be positive definite, and a small random disturbance w^α must be added to the unknown parameters α to insure controllability. Assuming that R, Q^x and Q^α are positive definite, the positive definiteness of the integrals of (6.9b) and (6.9c) will depend on the amount of "maneuvering" in the Φ^x and Φ^α matrices, as was observed for the special cases which were considered in Section 5.1.

Kalman (Lemma 16.9 and Lemma 16.10 [5]) has obtained the following bounds for P(t), under the assumption that the linear estimation problem is uniformly completely observable and uniformly completely controllable. The upper bound requires that P_0 be non-negative definite, and the lower bound requires that P_0 be positive definite. For $t \geq t_0 + \sigma$,

$$\left[W^{-1}(t-\sigma, t)+M(t-\sigma, t)\right]^{-1} \leq P(t) \leq M^{-1}(t-\sigma, t)+W(t-\sigma, t) \quad (6.10)$$

Property 6.3 The solution $P(t)$ $t \in [t_0, T]$, of the Riccati equation (6.7a) is symmetric and non-negative definite.

Demonstration Since R and Q are symmetric, $P(t)$ and its transpose $P^*(t)$ satisfy the same differential equation and, since P_0 is symmetric, $P(t_0)$ and $P^*(t_0)$ both satisfy the same initial conditions. Assuming that suitable bounds as in (6.7d) are obtainable, the uniqueness of the solution to (6.7a) implies that $P(t) = P^*(t)$ for $t \in [t_0, T]$. The non-negative definiteness of $P(t)$ is plausible if it is noted that the variance in the estimate of a costate y^* (see Theorem 1, [1]) is given by

$$E\left[(y^*, y(t)-\hat{y}(t))\right]^2 = \langle y^*, P(t)y^* \rangle \geq 0$$

6.4 The Necessary Conditions for the Optimization Problem

In this section the necessary conditions which the optimal control must satisfy are specified. The matrix form of the Hamiltonian and the adjoint equations have been previously stated in [11]. The proofs which require only straightforward matrix manipulations are included. To be specific, the problem formulated in section 6.1 (Problem I) is considered.

Property 6.4 If $\underline{\lambda}_0$ is an m x 1 adjoint vector which corresponds to the integral constraint $\int_I \underline{u}(t)dt = \underline{c}$, and Λ is a matrix of adjoint variables with elements λ_{ij} which correspond to the elements p_{ij} of the covariance matrix P, $i, j. = 1, \cdots, n$, then the Hamiltonian for this problem is given by

$$H = \langle \underline{\lambda}_0, \underline{u} \rangle + \text{Trace}\{[AP + PA^* - PM^*R^{-1}MP + Q]\Lambda^*\} \qquad (6.11)$$

Demonstration As in section 6.2, the matrix Riccati equation can be expanded as

$$\dot{p}_{ij}(t) = F_{ij}(P, u, t), \; i, j = 1, \cdots, n \qquad (6.12a)$$

and we define a further state vector by

$$\dot{\underline{x}}_0(t) = \underline{u}(t) \; . \qquad (6.12b)$$

Applying the Maximum Principle of Pontryagin, there results

$$\begin{aligned} H &= \langle \underline{\lambda}_0, \underline{u} \rangle + \sum_{i=1}^{n} \sum_{j=1}^{n} \dot{p}_{ij} \lambda_{ij} \\ &= \langle \underline{\lambda}_0, \underline{u} \rangle + \text{Trace}\left\{ \begin{bmatrix} \dot{p}_{11} \cdots \dot{p}_{1n} \\ \vdots \quad \vdots \\ \dot{p}_{n1} \cdots \dot{p}_{nn} \end{bmatrix} \begin{bmatrix} \lambda_{11} \cdots \lambda_{1n} \\ \vdots \quad \vdots \\ \lambda_{n1} \cdots \lambda_{nn} \end{bmatrix} \right\} \\ &= \langle \underline{\lambda}_0, \underline{u} \rangle + \text{Trace}\left[\left(\frac{d}{dt} P \right) \Lambda^* \right] \end{aligned}$$

and (6.11) follows when the right side of (6.2) is substituted for $\frac{d}{dt} P$.

Remark 6.1 If the assumptions of section 5.3 are made, then (6.11) may be written as equation (6.13c) below, where

$$\Lambda = \begin{bmatrix} \Lambda^{xx} & \Lambda^{x\alpha} \\ \Lambda^{\alpha x} & \Lambda^{\alpha\alpha} \end{bmatrix} \tag{6.13a}$$

and

$$\begin{aligned} \mathcal{P}^{xx} &= C^{\alpha} P^{\alpha x} + P^{x\alpha} C^{\alpha *} - r^{-1} P^{xx} P^{xx} + q \cdot I \\ \mathcal{P}^{x\alpha} &= C^{\alpha} P^{\alpha\alpha} - r^{-1} P^{xx} P^{x\alpha} \\ \mathcal{P}^{\alpha x} &= P^{\alpha\alpha} C^{\alpha *} - r^{-1} P^{\alpha x} P^{xx} \\ \mathcal{P}^{\alpha\alpha} &= -r^{-1} P^{\alpha x} P^{x\alpha}. \end{aligned} \tag{6.13b}$$

$$H = \lambda_o \mu + \text{Trace}[\mathcal{P}^{xx} \Lambda^{xx} + \mathcal{P}^{x\alpha} \Lambda^{\alpha x}] + \text{Trace}[\mathcal{P}^{\alpha x} \Lambda^{x\alpha} + \mathcal{P}^{\alpha\alpha} \Lambda^{\alpha x}] \tag{6.13c}$$

Property 6.5 Along the optimal trajectory, the adjoint variables satisfy the differential equations

$$\dot{\lambda}_o = \underline{0}, \tag{6.14a}$$

$$\dot{\Lambda} = -A^* \Lambda - \Lambda A + M^* R^{-1} M P \Lambda + \Lambda P (M^* R^{-1} M) \tag{6.14b}$$

with the end condition

$$\Lambda(T) = \frac{\partial}{\partial P(T)} \text{trace}[WP(T)] = W \tag{6.14c}$$

Demonstration Applying the maximum principle of Pontryagin, along the optimal trajectory gives

$$\frac{d}{dt}\underline{\lambda}_o = -\frac{\partial H}{\partial \underline{x}_o} = \underline{0}$$

and

$$\frac{d}{dt}\lambda_{ij} = -\frac{\partial H}{\partial p_{ij}}, \quad i,j = 1,\ldots,n.$$

The terms in the definition of H are considered one at a time.

$$\mathrm{Tr}[AP\Lambda^*] = \mathrm{Tr}\left\{A\begin{bmatrix} \sum p_{1j}\lambda_{1j} & \cdots & \sum p_{1j}\lambda_{nj} \\ \vdots & & \vdots \\ \sum p_{nj}\lambda_{ij} & \cdots & \sum p_{nj}\lambda_{nj} \end{bmatrix}\right\} =$$

$$= a_{11}\sum p_{1j}\lambda_{1j} + \cdots + a_{1n}\sum p_{nj}\lambda_{1j} +$$

$$+ a_{21}\sum p_{1j}\lambda_{ij} + \cdots + a_{2n}\sum p_{nj}\lambda_{2j} +$$

$$\cdots$$

$$+ a_{n1}\sum p_{1j}\lambda_{nj} + \cdots + a_{nn}\sum p_{nj}\lambda_{nj}.$$

Thus the matrix of partial derivatives

$$\frac{\partial(\mathrm{tr}[AP\Lambda^*])}{\partial P} = \left[\frac{\partial(\mathrm{tr}[AP\Lambda^*])}{\partial p_{ij}}\right] =$$

$$= \begin{bmatrix} a_{11}\lambda_{11} + \cdots + a_{n1}\lambda_{n1} & \cdots & a_{1n}\lambda_{1n} + \cdots + a_{n1}\lambda_{nn} \\ \vdots & & \vdots \\ a_{1n}\lambda_{11} + \cdots + a_{nn}\lambda_{n1} & \cdots & a_{1n}\lambda_{1n} + \cdots + a_{nn}\lambda_{nn} \end{bmatrix} = A^*\Lambda \qquad (6.15a)$$

Next we consider the term

$$\text{Trace}[PA^{*}\Lambda^{*}]=\text{Trace}\left\{P\begin{bmatrix} a_{j1}\lambda_{1j}\cdots & a_{j1}\lambda_{nj} \\ \vdots & \vdots \\ a_{jn}\lambda_{1j}\cdots & a_{jn}\lambda_{nj}\end{bmatrix}\right\} =$$

$$= p_{11}\sum a_{j1}\lambda_{1j} + \cdots + p_{1n}\sum a_{jn}\lambda_{1j} +$$

$$+ \cdots +$$

$$+ p_{n1}\sum a_{j1}\lambda_{nj} + \cdots + p_{nn}\sum a_{jn}\lambda_{nj}$$

Thus the matrix of partial derivatives becomes

$$\frac{\partial(\text{tr}[PA^{*}\Lambda^{*}])}{\partial p}=\begin{bmatrix} \sum a_{j1}\lambda_{1j} \cdots \sum a_{jn}\lambda_{1j} \\ \vdots \\ \sum a_{j1}\lambda_{nj} \cdots \sum a_{jn}\lambda_{nj}\end{bmatrix} = \Lambda A. \qquad (6.15b)$$

Next

$$\text{Trace}[PM^{*}R^{-1}MP\Lambda^{*}] = \text{Trace}\,[PM'P\,\Lambda^{*}]$$

$$= \text{Trace}\left\{PM'\begin{bmatrix} \sum p_{1j}\lambda_{1j} \cdots \sum p_{1j}\lambda_{nj} \\ \vdots \\ \sum p_{nj}\lambda_{1j} \cdots \sum p_{nj}\lambda_{nj}\end{bmatrix}\right\} =$$

$$= \begin{cases} \sum p_{1j}m_{j1}\sum p_{1j}\lambda_{1j} + \cdots + \sum p_{1j}m_{jn}\sum p_{nj}\lambda_{1j} + \\ \qquad + \cdots + \\ + \sum p_{nj}m_{j1}\sum p_{1j}\lambda_{1j} + \cdots + \sum p_{nj}m_{jn}\sum p_{nj}\lambda_{nj} \end{cases}$$

The matrix of partial derivatives becomes

$$\begin{bmatrix} m_{11}\sum p_{1j}\lambda_{1j} + \ldots + m_{1n}\sum p_{nj}\lambda_{1j} & \cdots & m_{n1}\sum p_{1j}\lambda_{1j} + \ldots + m_{nn}\sum p_{nj}\lambda_{1j} \\ \vdots & & \vdots \\ m_{11}\sum p_{1j}a_{nj} + \ldots + m_{1n}\sum p_{nj}\lambda_{nj} & \cdots & m_{n1}\sum p_{1j}\lambda_{nj} + \ldots + m_{nn}\sum p_{nj}\lambda_{nj} \end{bmatrix} +$$

$$+ \begin{bmatrix} \lambda_{11}\sum p_{1j}m_{j1} + \ldots + \lambda_{n1}\sum p_{nj}m_{j1} & \cdots & \lambda_{1n}\sum p_{1j}m_{j1} + \ldots + \lambda_{nn}\sum p_{nj}m_{j1} \\ \vdots & & \vdots \\ \lambda_{11}\sum p_{1j}m_{jn} + \ldots + \lambda_{n1}\sum p_{nj}m_{jn} & \cdots & \lambda_{1n}\sum p_{1j}m_{jn} + \ldots + \lambda_{nn}\sum p_{nj}m_{jn} \end{bmatrix} =$$

$$= \begin{bmatrix} \lambda_{11}\cdots\lambda_{1n} \\ \lambda_{n1}\cdots\lambda_{nn} \end{bmatrix} \begin{bmatrix} p_{11}\cdots p_{n1} \\ p_{1n}\cdots p_{nn} \end{bmatrix} \begin{bmatrix} m_{11}\cdots m_{n1} \\ m_{1n}\cdots m_{nn} \end{bmatrix} + \begin{bmatrix} m_{11}\cdots m_{n1} \\ m_{1n}\cdots m_{nn} \end{bmatrix} \begin{bmatrix} p_{11}\cdots p_{n1} \\ p_{1n}\cdots p_{nn} \end{bmatrix} \begin{bmatrix} \lambda_{11}\cdots\lambda_{1n} \\ \lambda_{n1}\cdots\lambda_{nn} \end{bmatrix} =$$

$$= \Lambda P^* M'^* + M'^* P^* \Lambda \,. \qquad (6.15c)$$

Using the results of (6.15a), (6.15b), (6.15c) and setting up the matrix of adjoint variables

$$\frac{d}{dt}\,\Lambda = -\left[\frac{\partial H}{\partial p_{ij}}\right]$$

yields equation (6.14b).

Since the "state" variables p_{ij} are free at $t=T$, the corresponding adjoint variables satisfy

$$\lambda_{ij}(T) = \frac{\partial}{\partial p_{ij}(T)} \left(\text{Trace}\left[WP(T)\right]\right)$$

along the optimal trajectory. This may be written as

$$\Lambda(T) = \left[\frac{\partial}{\partial p_{ij}(T)}\left(\sum w_{ij}\, p_{ij} + \cdots + \sum w_{nj}\, p_{jn}\right)\right] = W.$$

Property 6.6 There exists a solution of the adjoint equation (6.14b) which is unique and symmetric for $t \in I \equiv [\, t_0, T\,]$. Furthermore, if the weighting matrix W is positive definite, then this solution is positive definite for $t \in I$.

Demonstration The uniqueness result follows from the fact that $\Lambda(t)$ satisfies a linear homogeneous equation in Λ. Thus (6.14b) can be written as an $n^2 \times 1$ vector equation

$$\dot{\underline{\Lambda}} = G(t)\,\underline{\Lambda}, \quad \underline{\Lambda}(T) = W.$$

The elements of the $n^2 \times n^2$ matrix $G(t)$, which involve the optimal covariance elements $p_{ij}(t)$ and the control $\underline{u}(t)$, are integrable on $[\, t_0, T\,]$. Therefore Theorem 5.1, Appendix [8] implies that the solution $\Lambda(t)$, such that $\Lambda(T) = W$, exists and is unique. To show that the solution is symmetric, we let $B = M^* R^{-1} M$, and take the transpose of both sides of (6.14b). This gives

$$[\frac{d}{dt}\Lambda\ ]^* = \frac{d}{dt}\Lambda^* = -\Lambda^*A - A^*\Lambda^* + \Lambda^*PB + BP\Lambda^*,$$

and $\Lambda^*(T) = W^* = W$. Thus $\Lambda^*(t)$ and $\Lambda(t)$ satisfy the same differential equation. Since the solution to this equation is unique, it follows that $\Lambda^*(t) = \Lambda(t)$, for $t \in I$. The positive definiteness of $\Lambda(t)$ is given by Property 7.1, Chapter VII

Remark 6.2 For the system specified in section 5.3, equations (6.14b) and (6.14c) become

$$\frac{d}{dt}\Lambda^{xx} = r^{-1}(P^{xx}\Lambda^{xx} + \Lambda^{xx}P^{xx} + P^{x\alpha}\Lambda^{\alpha x} + \Lambda^{x\alpha}P^{\alpha x});\ \Lambda^{xx}(T)=0 \quad (6.16a)$$

$$\frac{d}{dt}\Lambda^{x\alpha} = -\Lambda^{xx}A_i + r^{-1}(P^{xx}\Lambda^{x\alpha} + P^{x\alpha}\Lambda^{\alpha\alpha});\ \Lambda^{x\alpha}(T) = 0 \quad (6.16b)$$

$$\frac{d}{dt}\Lambda^{\alpha\alpha} = -(A_i^*\Lambda^{\alpha\alpha} + \Lambda^{\alpha x}A_i)\ ;\ \Lambda^{\alpha\alpha}(T) = W_\alpha \quad (6.16c)$$

Property 6.7 For the cost functional

$$J(\underline{u}) = \text{Trace } W P(T;\underline{u}), \quad (6.17a)$$

an optimal control $\underline{\bar{u}}$ satisfies the condition

$$\begin{aligned} &\langle \underline{\lambda}_o, \underline{\bar{u}} \rangle + \text{Trace}\,[A(\underline{\bar{u}})P\Lambda^* + PA^*(\underline{\bar{u}})\Lambda^*] \leq \\ &\leq \langle \underline{\lambda}_o, \underline{u} \rangle + \text{Trace}[A(\underline{u})P\Lambda^* + PA^*(\underline{u})\Lambda^*] \end{aligned} \quad (6.17b)$$

for all $\underline{u} \in \mathcal{U}$, where $\mathcal{U}$ is specified by

$$\mathcal{U} = \{\underline{u} : \underline{u}(t) \in U(t),\ t \in I,\ \int_I \underline{u}\,dt = \underline{c}\}\ . \quad (6.17c)$$

In particular, if $U(t)$ is a "cube", that is

$$U(t) = \{u(t) : |u_i(t)| \leq M_i \text{ for } t \in I,\ i = 1, \cdots, m\} \quad (6.17d)$$

then the optimal controls are "bang-bang", that is,

$$\bar{u}_i(t) = \pm M_i \text{ for } t \in I \tag{6.17e}$$

This result, of course, assumes that singular controls do not arize.

Demonstration Equation (6.17b) is a direct result of the definition of the Hamiltonian H and the maximum principle since the control $\underline{u}$ enters the matrix A only. To obtain the "bang-bang" result (6.17d), it is recalled that $\underline{u}$ enters the matrix A in a linear fashion. Thus that portion of the left sides of (6.17b) which involve $\underline{u}$ can be written as

$$\sum_{i=1}^{m} \bar{u}_i \left(\lambda_o^i + h_i(P,\Lambda, t)\right) \tag{6.17f}$$

Thus to minimize the Hamiltonian in the case U is defined by (6.17d), we choose

$$\bar{u}_i = -M_i \text{ sign } \left(\lambda_o^i + h_i(P,\Lambda)\right) \tag{6.17g}$$

This implies $\bar{u}_i(t) = \pm M_i$, provided that $\lambda_o^i + h_i(P,\Lambda) \neq 0$. If the latter condition results for a finite time interval, then $\bar{u}_i(t)$ is indeterminate.

Remark 6.3 In the case of Configurations I and II,

$$A P\Lambda^* + P A^* \Lambda^* = \left\{ \begin{bmatrix} 0 & A_i \\ 0 & 0 \end{bmatrix} \begin{bmatrix} P^{xx} & P^{x\alpha} \\ P^{\alpha x} & P^{\alpha\alpha} \end{bmatrix} + \begin{bmatrix} P^{xx} & P^{x\alpha} \\ P^{\alpha x} & P^{\alpha\alpha} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ A_i^* & 0 \end{bmatrix} \right\} \begin{bmatrix} \Lambda^{xx} & \Lambda^{x\alpha} \\ \Lambda^{\alpha x} & \Lambda^{\alpha\alpha} \end{bmatrix}$$

which results in a control law

$$F(t) = -M \operatorname{sign}\left[\lambda_o + \operatorname{Trace}(A_i P^{\alpha x} + P^{x\alpha} A_i^*)\Lambda^{xx} + \operatorname{Trace}(A_i P^{\alpha\alpha}\Lambda^{\alpha x} + P^{\alpha\alpha} A_i^* \Lambda^{\alpha\alpha}\right] \tag{6.18}$$

where $i = I$ for Configuration I and II for Configuration II.

Remark 6.4 By considering certain monotone properties of the elements of Λ and P, it is possible to obtain upper bounds on the number of switches the optimal control will have. In particular, it is clear that the diagonal elements of $P^{\alpha\alpha}$ are non-increasing. If certain assumptions about the noise covariance matrices Q and R can be made as, for example, they are also non-increasing, then it is plausible that $P^{x\alpha}$ and P^{xx} have similar monotone properties.

6.5 Unconstrained Controls

If it is assumed that the controls do not have pointwise constraints and the cost functional is of the form (6.19c), with the integral constraint $\int_I \underline{u}(t)\,dt = \underline{c}$, then the optimal control is specified by (6.19b), where the components of $\underline{h}$, and h_i, are determined by (6.19c). In order to obtain such smooth controls and at the same time insure that the control components $u_i(t)$ are not too large, it might be possible to adjust the weighting matrix $C(t)$.

$$J_2(\underline{u}) = \frac{1}{2}\int_{T_1} \|\underline{u}(t)\|_C^2 + \text{Trace}\,[\,W P(T\,;\underline{u})] \qquad (6.19a)$$

$$\bar{\underline{u}}(t) = -C^{-1}(t)[\,\underline{\lambda}_o + \underline{h}\,(P,\Lambda,t)] \qquad (6.19b)$$

$$\frac{\partial}{\partial u_i}\left\{\text{Trace}\left[\,AP\Lambda + PA^*\Lambda\right]\right\} \qquad (6.19c)$$

In particular, for Configurations I and II the optimal control is specified by

$$\bar{\mu}(t) = -\frac{1}{c(t)}\left\{\lambda_o + \text{tr}\left[(A_i P^{\alpha x} + P^{x\alpha}A_i^*)\Lambda^{xx} + (A_i P^{\alpha\alpha}\Lambda^{\alpha x} + P^{\alpha\alpha}A_i^*\Lambda^{x\alpha})\right]\right\} \qquad (6.19d)$$

If pointwise constraints are also included for these special configurations then the optimal control is specified by (6.19e) where μ_{min} is the right side of (6.19d). This result follows from the fact that $\partial^2 H(u)/\partial u^2 = c > 0$.

$$\bar{\mu}(t) = \begin{cases} -M_1 & \text{if } \mu_{min}(t) \leq -M_1 \\ \mu_{min}(t) & \text{if } \mu_{min}(t) \leq M_1 \\ +M_1 & \text{if } \mu_{min}(t) \geq M_1 \end{cases} \qquad (6.19e)$$

Property 6.8 For the unconstrained control problem formulated above with cost functional (6.19a) and with no integral constraint and with W positive definite, there exists an admissible optimal control $\bar{\underline{u}}$ such that $J_2(\underline{u})$ achieves a global minimum.

Demonstration Choosing the control $\underline{u}(t) \equiv \underline{0}$, the corresponding response is $P(t;\underline{o})$ and the cost is $J(\underline{o}) = \text{Trace}\,[\,WP(T\,;\underline{o})\,]$. A finite bound for $\|P(t,\underline{o})\|_\infty \equiv \sup_{ij} |p_{ij}(t)|$ can always be

obtained as discussed in section 6.3 or in the proof of Property 6.1. It is next asserted that all other controls $\underline{u}_i$ which yield a cost such that

$$J(\underline{u}_i) \leq J(\underline{o})$$

will satisfy $\|P(t;\underline{u}_i)\|_\infty \leq \|P(t;\underline{o})\|_\infty$ for $t \in I$. This is due to the fact that at each instant of time $t' \in I$ there is the same amount of control available when either $\underline{u}(t) = \underline{u}_i(t)$ or $\underline{u}(t) = \underline{o}$ were used for $t \in [t_o, t']$. Thus if the trace $P(t;u_i) >$ trace $P(t;\underline{o})$ at $t=t'$, the control

$$\underline{u}(t) = \begin{cases} \underline{o} & t_o \leq t \leq t' \\ \underline{u}_i(t) & t' \leq t \leq T \end{cases}$$

would yield a smaller cost than $\underline{u}_i$ since $W > 0$. Since the norms in E_n are topologically equivalent, the assertion follows. It is remarked that this result is not necessarily true if the integral constraint on $\underline{u}$ is in force, or if $W \geq 0$. Further, since all the controls $\underline{u}_i$ yield a finite cost, we have

$$J(\underline{u}_i) \geq \int_I \|\underline{u}_i(t)\|_C^2 \, dt \geq c \int_I \|\underline{u}_i(t)\|^2 \, dt, \quad 0 < c < \infty . \qquad (6.20a)$$

If there is only a finite number of controls with the above property, then the existence of an optimal control is immediate. Otherwise, a subsequence $\{\underline{u}_{i'}\}$ of these controls can be selected such that $J\{\underline{u}_{i'}\}$ tends monotonically towards the infimum j of J. Further, since the unit sphere in L_2 is weakly

sequentially compact, a further subsequence can be selected such that

$$\underline{u}_{i'}(t) \xrightarrow{w} \bar{\underline{u}}(t) \text{ with } \bar{\underline{u}} \in L_2([t_0, T])$$

The responses $\{\underline{P}_{i'}\}$ to the controls $\{\underline{u}_{i'}\}$ will be equibounded, and we can show that they are also equicontinuous. Indeed, for two times $t_0 \le t'$, $t'' \le T$, we have

$$\|\underline{P}_{i'}(t') - \underline{P}_{i'}(t'')\|_\infty \le \int_{t'}^{t''} \left\{ \|\underline{G}(\underline{P}_{i'}(t), t)\|_\infty + \|\underline{H}(\underline{P}_{i'}(t), t)\underline{u}_{i'}(t)\|_\infty \right\} dt$$

Since the partial derivatives $\partial \dot{p}_{ij}/\partial p_{k\ell}$ are continuous for $i, j, k, \ell = 1, \cdots, n$, $\underline{P} \in E_n$ and $\underline{u} \in E_m$, the above inequality becomes

$$\begin{aligned} &\le c_1 \, |t''-t'| + c_2 \int_{t'}^{t''} \|\underline{u}_{i'}(t)\|_1 \, dt \\ &\le c_1 \, |t''-t'| + c_2 \, |t''-t'|^{1/2} \left(\int_{t'}^{t''} \|\underline{u}_{i'}(t)\|_2^2 \, dt \right) \end{aligned} \tag{6.20b}$$

However, (6.20a) implies that for i' sufficiently large,

$$c \int_I \|\underline{u}_{i'}(t)\|_2^2 \, dt \le j + \epsilon, \quad 0 \le \epsilon < \infty . \tag{6.20c}$$

Inequalities (6.20b) and (6.20c) give the required result

$$\|\underline{P}_{i'}(t'') - \underline{P}_{i'}(t')\|_\infty \le c_1 \, |t''-t'| + c_2' \, |t''-t'|^{1/2} \tag{6.20d}$$

Since the sequence $\{\underline{P}_{i'}\}$ is equibounded and equicontinuous, Ascoli's Theorem insures the existence of a subsequence $\{\underline{P}_{i''}\}$ such that

$$\underline{P}_{i''}(t) \longrightarrow \bar{\underline{P}}(t) \text{ uniformly for each } t \in I.$$

The fact that the $\underline{u}_{i'}$ enter the dynamical equations linearly is

then used to take weak limits and show that $\bar{P}$ is the response to $\underline{\bar{u}}$, the weak limit of $\{\underline{u}_i\}$. The convexity of $\|u(t)\|_C^2$ will imply that the cost functional is weakly lower semicontinuous. Thus $J(\underline{\bar{u}}) = j$, and $\underline{\bar{u}}$ is the required optimal control.

6.6 An Example

At this point it is instructive to consider the system below as a simple illustration of the above ideas.

$$\dot{x} = ax + u\alpha + w^x, \qquad \operatorname{cov}(w^x) = q^x ;$$

$$\dot{\alpha} = 0 + w^\alpha, \qquad \operatorname{cov}(w^\alpha) = q^\alpha ;$$

$$y = x + v, \qquad \operatorname{cov}(v) = r ;$$

Thus

$$A = \begin{bmatrix} a & u \\ 0 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} q^x & 0 \\ 0 & q^\alpha \end{bmatrix}, \quad R = r, \quad M = [1, 0],$$

and the matrix Riccati equation

$$\dot{P} = AP + PA^* - PM^* R^{-1} MP + Q$$

yields the state equations

$$\begin{cases} \dot{p}_{11} = 2(ap_{11} + up_{21}) - r^{-1} p_{11}^2 + q^x, & p_{11}(0) \text{ given;} \\ \dot{p}_{12} = ap_{12} + up_{22} - r^{-1} p_{11} p_{12}, & p_{12}(0) = 0 ; \\ \dot{p}_{22} = -r^{-1} p_{12}^2 + q^\alpha, & p_{22}(0) \text{ given.} \end{cases}$$

Assuming the integral constraint $\int_0^T u(t)\,dt = 0$, we define the additional state equation

$$\dot{x}_o = u, \qquad x_o(0) = x_o(T) = 0$$

The necessary conditions specified by

$$\dot{\Lambda} = -A^* \Lambda - \Lambda A + M^* R^{-1} M P \Lambda + \Lambda P^* (M^* R^{-1} M)^*$$

yields the adjoint equations

$$\begin{cases} \dot{\lambda}_{11} = 2(r^{-1} p_{11} - a)\lambda_{11} + 2 r^{-1} p_{12} \lambda_{12} \\ \dot{\lambda}_{12} = -u \lambda_{11} + (r^{-1} p_{11} - a)\lambda_{12} + r^{-1} p_{12} \lambda_{22} \\ \dot{\lambda}_{22} = -2 u \lambda_{12} \\ \dot{\lambda}_o = 0 \end{cases}$$

Next the constraint $|u| \leq M_1$, and the two cost functionals

$$J_1(u) = c^{\alpha} p_{22}(T;u) + c^{x} p_{11}(T;u)$$

$$J_2(u) = c^{\alpha} p_{22}(T;u) + c^{u} \int_0^T u^2(t)\, dt$$

are assumed. For J_1 and J_2 the boundary conditions for λ_{ij} are, respectively

$$\lambda_{11}(T) = c^{x} \qquad \lambda_{11}(T) = 0$$

$$\lambda_{12}(T) = 0 \quad \text{and} \quad \lambda_{12}(T) = 0$$

$$\lambda_{22}(T) = c^{\alpha} \qquad \lambda_{22}(T) = c^{\alpha}$$

and the initial values of the λ_{ij} are unknown. For J_1, the Hamiltonian is

$$H_1(u) = \lambda_o u + \text{Trace}[(AP + PA^* - PM^* R^{-1} MP + Q)\Lambda^*]$$

$$= \lambda_o u + \lambda_{11}(2(ap_{11} + up_{21}) - r^{-1} p_{11}^2 + q^{x}) +$$

$$+ 2\lambda_{12}(ap_{12} + up_{22} - r^{-1} p_{11} p_{12}) + \lambda_{22}(q^{\alpha} - r^{-1} p_{12}^2)$$

which means that the optimal control is specified by

$$u(t) = -M_1 \operatorname{sign} (\lambda_c + 2\lambda_{11}p_{21} + 2\lambda_{12}p_{22}) .$$

Since $H_1(t) = H(T) - \int_t^T (\partial H/\partial t)\, dt$, and $\partial H/\partial t = 0$, we can set $H(0) = H(T)$ to obtain

$$-M_1|\lambda_0| + c^{\alpha}(-r^{-1}p_{12}^2(T) + q^{\alpha} = -M|\lambda_0 + 2\lambda_{12}(0)p_{22}(0)| +$$
$$+ \lambda_{11}(0)(2ap_{11}(o) - r^{-1}p_{11}^2(0) + q^{\alpha}) + \lambda_{22}(0) + q^{\alpha} .$$

This equation can be used to eliminate one of the unknown boundary values by considering the various possibilities for the signs of λ_0 and $\lambda_0 + 2\lambda_{12}(0)p_{22}(0)$.

Numerical results show that the optimal control is constant if there is no integral constraint, and switches once if the integral constraint is in force. For the cost functional J_2, minimizing the Hamiltonian implies that it is necessary to minimize

$$h(u) \equiv (\lambda_c + 2\lambda_{11}p_{21} + 2\lambda_{12}p_{22}) + c^u u^2$$

Since $\partial^2 h(u)/\partial u^2 = c^u > 0$, we set

$$u_{min} \equiv -(\lambda_0 + 2\lambda_{11}p_{21} + \lambda_{12}p_{22})/c^u,$$

and the optimal control is specified by

$$u(t) = -M_1 \quad \text{if} \quad u_{min} \leq -M_1$$

$$u(t) = u_{min} \quad \text{if} \quad |u_{min}| \leq M_1$$

$$u(t) = +M_1 \quad \text{if} \quad u_{min} \geq M_1$$

If u is unbounded, then $u(t) = u_{min}(t)$.

CHAPTER VII

COMPUTATIONAL ASPECTS FOR COMPUTING THE OPTIMAL CONTROLS

In order to compute the optimal controls using the necessary conditions of the previous chapter, it is necessary to solve the two-point boundary value problem associated with the matrix Riccati equation and the matrix adjoint equation. The initial condition for the covariance matrix P is specified by P_0, and, generally, the final values for the matrix of adjoint variables Λ is specified by Λ_T, which depends on the particular cost functional which is being minimized. Many people have been concerned with the numerical solution of similar two point boundary value problems. The success of a particular scheme depends on having considerable insight about the solutions of the particular problem being solved, since a fairly accurate guess of the unknown adjoint variables Λ_0 is required. In fact, many published results which show successful iteration schemes for choosing the unknown boundary condition for reasonably complicated problems, start with initial guesses in which the cost functional is quite close to being a minimum. The nature of the problem considered here is such that considerable computational effort can be justified for the numerical solution of a particular

identification problem due to the practical interest in improved identification. A particular example in which improved identification of IMU error parameters is very profitable occurs in the flight testing of inertial platforms for purposes of obtaining the instrument errors under actual operating conditions. In these cases it is necessary to "design the experiment", that is, select a nominal trajectory for the flight test vehicle so as to obtain as much information about the error parameters as possible for a particular flight. The material presented in the previous chapters offers a systematic way to maximize this information. The same remarks can be made in regards to inflight alignment and calibration of inertial measurement units, and for the flight testing of other devices in which it is desirable to identify error parameters.

Due to the complicated nature of the present optimization problem, we are primarily interested in computation schemes in which it is possible to observe certain iterations of the computation and then choose new boundary values based on these observations. The possibility of using algorithms which are based on a direct minimization of the cost functionals involved are not considered, although such algorithms are suitable topics for further investigations. In particular, the ϵ-method which has recently been studied for optimal control problems [16] could be considered.

7.1 Iteration on the Initial Adjoint Variables, Λ_0

For this scheme a cost functional involving Λ_T and P_T is specified, for example, as

$$I(\Lambda_0) = (\tilde{\underline{P}}_T - \underline{P}_T)^* W^P (\tilde{\underline{P}}_T - \underline{P}_T) + (\tilde{\underline{\Lambda}}_T - \underline{\Lambda}_T) W^\lambda (\tilde{\underline{\Lambda}}_T - \underline{\Lambda}_T) \qquad (7.1)$$

where the tilda denotes a final value of P or of Λ resulting from an initial guess on the adjoint variables Λ_0, and W^P and W^λ are non-negative weighting matrices. If an integral constraint on the controls is included, say of the form $\int_I \underline{u}(t)\,dt = \underline{c}$, then we let $\tilde{\underline{x}}_0(T) = \int_0^T \tilde{\underline{u}}(t)\,dt$, and add a term of the form $\langle(\tilde{\underline{x}}_0(T) - \underline{c}), W^u(\tilde{\underline{x}}_0(T) - \underline{c})\rangle$ to the functional (7.1). Associated with this cost functional are the differential equations

$$\dot{P} = AP + PA^* - PBP + Q, \qquad P(t_0) = P_0 \text{ given;} \qquad (7.2)$$

$$\dot{\Lambda} = -A^*\Lambda - \Lambda A + BP\Lambda + \Lambda PB, \quad \Lambda(T) = \Lambda_T \text{ may be given.} \qquad (7.3)$$

and the control equation (which results from minimizing the Hamiltonian)

$$\underline{u}(t) = \underline{h}(P, \Lambda, t) \qquad (7.4)$$

The idea in the minimization is to choose a sequence $\{\Lambda_0^{(i)}\}$ of initial adjoint variables so that $I(\Lambda_0^{(i)})$ approaches zero after a reasonable number of iterations.

A straightforward method for choosing the successive iterations is to form $\partial I(\Lambda_o)/\partial \underline{\Lambda}_o$ and then proceed in the direction of steepest descent. Since $I(\Lambda_o) = I\Big(P_T(\Lambda_o), \Lambda_T(\Lambda_o)\Big)$, we have

$$\frac{\partial I}{\partial \underline{\Lambda}_o} = \frac{\partial I}{\partial \underline{P}_T} \cdot \frac{\partial \underline{P}_T}{\partial \underline{\Lambda}_o} + \frac{\partial I}{\partial \underline{\Lambda}_T} \frac{\partial \underline{\Lambda}_T}{\partial \underline{\Lambda}_o} = 2W^P(\tilde{\underline{P}}_T - \underline{P}_T)\frac{\partial \underline{P}_T}{\partial \underline{\Lambda}_o} + 2W^\lambda(\tilde{\underline{\Lambda}}_T - \underline{\Lambda}_T)\frac{\partial \underline{\Lambda}_T}{\partial \underline{\Lambda}_o} \quad (7.5)$$

The partial derivatives can be approximated by $\Delta \underline{P}_T/\Delta \underline{\Lambda}_o$ and $\Delta \underline{\Lambda}_T/\Delta \underline{\Lambda}_o$, respectively, where $\Delta \underline{P}_T$ and $\Delta \underline{\Lambda}_T$ denotes the changes in the elements of P_T and Λ_T due to incremental changes in the initial values of Λ. The approximate derivatives would be obtained by solving (7.2), (7.3), (7.4) as initial value problems $n(n+1)/2$ times. If all the elements of P_T are unspecified, then $W^P \equiv 0$ and the resulting computations are considerably reduced.

The main idea in the steepest descent approach is as follows:

(1) Make an initial guess $\Lambda_o^{(1)}$

(2) Compute $\partial I/\partial \underline{\Lambda}_o \big|_{\Lambda_0^{(1)}} \simeq \Delta I/\Delta \underline{\Lambda}_0^{(1)} \equiv \delta \underline{\Lambda}_0^{(i)}$

(3) Let $\underline{\Lambda}_0^{(2)} = \underline{\Lambda}_0^{(1)} + k \cdot \delta \underline{\Lambda}_0^{(1)}$, where $k < 0$

(4) If $I(\Lambda_0^{(2)}) < I(\Lambda_0^{(1)})$, return to step (1) and continue this process until $I(\Lambda_0^{(k)})$ is sufficiently close to zero.

There are several methods available for choosing the "gain" k. The value of k should be small enough so that the minimization proceeds in the right direction and yet k must be sufficiently

large so that the minimization does not require too many iterations. One method for choosing k might be to specify a fixed percent change in the cost for such iteration, so that for η fixed, $\Delta I^{(i)} = -\eta I^{(i-1)}$. This results in $k^{(i)} = -\eta I^{(i-1)} / \|\partial I^{(i-1)} / \partial \underline{\Lambda}_o^{(i)}\|^2$. Another method is to choose k such that $I(\Lambda_o + k\Lambda_f)$ is minimized with respect to k where Λ_f is a prespecified matrix. This results in a k specified by $-\|\underline{\Lambda}_f\|^2 / \langle \underline{\Lambda}_f, \text{grad } I(\Lambda_f) \rangle$. If it is practical to be in the computation "loop", then it is probably more satisfactory to select k by observing the resulting changes in $I^{(i)}$. Thus if $I(\Lambda_o^{(i+1)}) \geq I(\Lambda_o^{(i)})$, then k is too large, and if $I^{(i+1)}$ is nearly equal to $I^{(i)}$, then k could be increased.

7.2 Iteration on the Final Covariance Matrix, P(T)

For particular problems, it might be more efficient to iterate on the final covariance matrix P_T instead of on the initial adjoint variables Λ_o. The iteration procedure is essentially as in the previous section, except that the adjoint and covariance equations are solved with the time variable reversed and the iteration is with P_T. More specifically, the cost functional (7.1) becomes

$$I(P_T) = \|\underline{\tilde{P}}_o - \underline{P}_o\|_{WP} + \|\underline{\tilde{\Lambda}}_o - \underline{\Lambda}_o\|_{W\lambda} \qquad (7.6)$$

and we let $\tau = T - t$ in equations (7.2), (7.3) and (7.4). This method has the advantage that the backwards covariance equation might be less "unstable" than the forward adjoint equation.

7.3 Selection of Initial Values, Λ_0 (or P_T)

Since the adjoint equation is highly unstable it may be impossible to obtain convergence in the above iteration schemes unless good initial choices of Λ_0 (or P_T) can be made. Firstly, using results of W.T. Reid [15] we show the

Property 7.1 If the cost functional is of the form trace $[WP(T;\underline{u})]$, then the solution $\Lambda(t)$ of the adjoint equation will be positive definite if W is positive definite.

Demonstration: W.T. Reid [15] considered the matrix differential equation

$$\dot{T} = H(t)\,T + TK(t) \tag{7.7}$$

and showed that the general solution is of the form $T(t) = U(t)CV(t)$, where $U(t)$ is a fundamental solution of $\dot{U} = HU$ and $V(t)$ is a fundamental solution of $\dot{V} = VK$, and C is an arbitrary constant matrix. The elements of H and K are assumed to be continuous on the interval $I = [t_0, T]$. If the elements of these matrices

are assumed only to be integrable, then the solution to (7.7) is a matrix whose elements are absolutely continuous and (7.7) is satisfied almost everywhere in I.

Under the assumption that $H = K^*$, it is clear that if $U(t)$ is a fundamental solution of $\dot{U} = KU$, then $U^*(t)$ is a fundamental solution of $\dot{V} = VK^*$, and the solution of (7.2) under this assumption will be $T(t) = U(t)\, C\, U^*(t)$. Now the adjoint equation can be written as

$$\dot{\Lambda} = K^*\Lambda + \Lambda K, \text{ where } K \equiv -A + (PM^*R^{-1}M)^* \qquad (7.8)$$

Using the usual notation $\Phi(t,T)$ for the fundamental solution of $\dot{V} = KV$, the solution of (7.8) is $\Lambda(t) = \Phi(t,T)\, C\Phi^*(t,T)$. Since $\Lambda(T) = W$, and W is positive definite by hypothesis, the solution is $\Lambda(t) = \Phi(t,T)\, W\Phi^*(t,T)$. This solution is positive definite since $\Phi^{-1}(t,T)$ always exists.

Remark This property could be of practical value in the selection of the initial values Λ_0. If the solution fails to remain positive definite for a particular positive definite choice of Λ_0, then a new choice can be made without necessarily completing the solution of the equations.

Since the optimal control can be assumed bang-bang in most instances, it is feasible to test the effect of various control functions upon the covariance functions in an open loop manner. If, for example, useful upper and lower bounds on the number of switchings were available, the search procedure would be greatly simplified. For the example introduced in Section 6. the effect of the number of switchings is shown in figure 7.1. In this example at least one switch is required because of the integral constraint $\int_I u(t)\,dt = 0$. These results show that a control which switches once at $t = 0.5$ seconds seems to be the optimal. Once several controls have been tried, the one which gives the best results could be used to solve the adjoint equation backwards (or else give P_T directly) to obtain a good initial guess $\Lambda_0^{(i)}$.

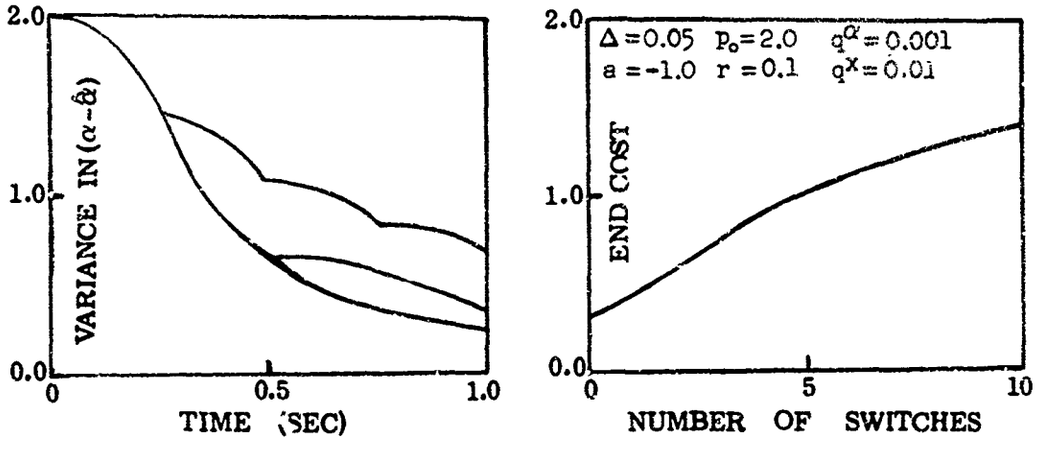


Figure 7.1 Effect of Switching on the Estimation of α

7.4 Discrete Form of the Covariance and Adjoint Equations

For computational purposes it may be more efficient to work with the discrete form of the covariance and adjoint equations. Actually, in some applications it is necessary to work with discrete state and observation equations. The discrete form of the matrix Riccati equation is given by

$$P_{k+1} = P'_{k+1} - P'_{k+1} M^*_{k+1} (M_{k+1} P'_{k+1} M^*_{k+1} + R_{k+1})^{-1} M_{k+1} P'_{k+1} \tag{7.9a}$$

$$P'_{k+1} = \Phi_{k+1,k} P_k \Phi^*_{k+1,k} + Q_k \tag{7.9b}$$

If $F_k = F_k(P_k, u_k)$ is defined by $P_{k+1} - P_k$, then, with the integral constraint on $\underline{u}$, the Hamiltonian becomes

$$H = \langle \underline{\lambda}_0, \underline{u}_k \rangle + \text{Trace}\,[F_k \Lambda^*_k] \tag{7.10}$$

and the adjoint variables will satisfy the equation

$$\Lambda_{k+1} - \Lambda_k = -\partial H / \partial P_k \tag{7.11}$$

along the optimal trajectory. The optimal control $\underline{u}_k$ is obtained by minimizing H. When it is possible to evaluate the matrix $\partial H / \partial P_k$, the adjoint equation will now be in discrete form. It is noted that in minimizing H_k, a neat result for the optimal $\underline{u}_k$ is

not obtained as in the continuous case because of the inverse matrix operations involved in equation (7.9). A good approximation which might be considered is to minimize the continuous version of the Hamiltonian for the selection of $\underline{u}_k$.

For the example presented in section 6.6, the covariance equation satisfies

$$P_{k+1} - P_k = \begin{bmatrix} r\pi\pi' - c & r\eta\pi' - d \\ r\eta\pi - d & q^{\alpha} - \eta^2\pi' \end{bmatrix} \tag{7.12a}$$

and the adjoint equation satisfies

$$\underline{\Delta}_{k+1} - \underline{\Delta}_k = -\begin{bmatrix} -1+(ra\pi')^2 & -2r\eta a^2\pi'^2 & (\eta a\pi')^2 \\ 2abr^2\pi'^2 & (2abr(\pi+r-\eta)\pi'^2-2) & -2a\eta\pi'^2(\pi+r-b\eta) \\ (rb\pi')^2 & 2rb\pi'^2(\pi+r-b\eta) & -b\eta\pi'^2(2(\pi+r)-b\eta) \end{bmatrix}\Delta_k \tag{7.12b}$$

In these equations the variables are defined as follows:

$$\begin{aligned} &a = e^{a\Delta} \qquad b = \tfrac{1}{a}[e^{a\Delta}-1]u_k \qquad ; \\ &c = P_{11}(k) \quad d = P_{12}(k) \qquad e = P_{22}(k) \; ; \\ &\pi = a\{ca+db)+b(da+cb)+q^{\chi} \qquad ; \\ &\eta = ad+be \qquad \pi' = 1/\pi + r \; . \end{aligned} \tag{7.12c}$$

For this example, with cost functional $J_1(u)$ of section 6.6, the iteration scheme is thus to choose the sequence $\{\underline{\lambda}^{(i)}\} \equiv \{\lambda_o^i, \lambda_{11}^{(i)}(o), \lambda_{12}^{(i)}(o), \lambda_{22}^{i}(o)\}^*$ such that $I(\lambda_o^{(i)})$ tends to zero, where

$$I(\underline{\lambda}_o^{(i)}) = (\tilde{\lambda}_{11}^{i}(T) - c^{x})^2 + (\tilde{\lambda}_{12}^{i}(T))^2 + (\tilde{\lambda}_{22}^{i}(T) - c^{\alpha})^2 + w^{x_o}(\tilde{x}_o(T))^2 \qquad (7.12d)$$

For each i, the vector of partial derivatives $\partial I^{(i)}(\Delta_o^{(i)})/\partial \underline{\lambda}_{(o)}^{(i)}$ was obtained by changing the initial adjoint vector $\underline{\lambda}_o^{(i)}$ by $\Delta\underline{\lambda}_o^{(i)}$ as shown in the table below, and then resolving equations (7.12a), (7.12b) and computing the optimal control from (7.12e), four times.

$$u^i(t) = -M \text{ sign } (\lambda_o^i + 2\lambda_{11}^i p_{21}^i + 2\lambda_{12}^i p_{22}^i), \qquad (7.12e)$$

Case	$\Delta\lambda_o$	$\Delta\lambda_{11}$	$\Delta\lambda_{12}$	$\Delta\lambda_2$	$\Delta I/\Delta\lambda_{ij}^i$
1	X	O	O	O	*
2	O	X	O	O	*
3	O	O	X	O	*
4	O	O	O	X	*

7.5 Future Problem Areas

The material presented in this study suggests numerous areas in which both practical and theoretical problems can be found. A few of the problems which the author feels might be fruitful areas for future research are mentioned in this section. This collection of problems is by no means complete, and, depending on one's particular practical or theoretical interests, additional problem areas would be formulated.

The techniques outlined in Chapter II might be used to study gravity errors and systems in which radar is used to obtain observation data. In the case of strapped-down IMU's, the elements of the

computed transformation matrix might be used to obtain additional observational data. As mentioned in Chapter II, angular velocity observations could also be used to improve the identification of the error parameters. Additional work (both experimental and analytical) must be directed towards obtaining realistic probabilistic descriptions of the environmental and observational random disturbances. To supplement the experimental data presented in Chapter IV additional specific error parameter configurations, such as configurations which would include accelerometer bias, scale-factor, and misalignment angles, should be considered to obtain additional insight into the trajectory optimization problem. Additional work on the "no-noise" properties presented in Chapter V would also be desirable.

A meaningful area of research in regards to control theory applications would be to study the properties of the two-point boundary-value problem associated with the matrix Riccati equation and the adjoint equation. Theorems on the number of switches the optimal control makes, whether or not singular controls result, and whether or not the optimal controls are unique, would be of great value in computing the optimal controls. Numerical experimentation with direct and indirect methods for computing the optimal controls should be made. Indirect methods (that is, in terms of

the adjoint equations) have been presented in Chapter VII. Although the resulting optimal controls are open-loop, these are of practical interest because the maneuvering associated with the parameter identification problems of flight-testing inertial components, and inflight alignment and calibration of IMU's can be precalculated.

Additional theoretical control problems arize when the control functions enter the dynamical equations in the complicated fashion suggested in section 5.2 for Configuration III. In this case (and in other cases in which mass-unbalance terms are included) the product of the control and its integral appear in the dynamical equations. Finally, there are systems in which the identification problem requires that a linearized minimum variance estimator be used. In this case a stochastic optimization problem must be considered, or else Monte Carlo studies would need to be made to compare the covariance which is assumed to be a solution of the Riccati equation with the true covariance.

REFERENCES

1. Kalman, R.E. and Bucy, R.S., "New Results in Linear Filtering and Prediction Theory", 1961 Trans. A.S.M.E. J. Basic Engineering 83D, pp. 95-108.

2. Pontryagin, L.S., Boltyanski, V.G., Gamkrelidze, R.V., and Mischenko, E.F., The Mathematical Theory of Optimal Processes, New York, Wiley, Interscience Publishers, 1962.

3. Baziw, J., "Inflight Alignment and Calibration of Master-Slave IMU Systems", TRW Systems Technical Report 07784-6134-R000, Redondo Beach, California.

4. Sorenson, H.J., "On the Error Behavior in Linear Minimum Variance Problems", IEEE Transactions on Automatic Control, Vol. AC-12, No. 5, Oct. 1967, pp. 557-562.

5. Kalman, R.E., "New Methods and Results in Linear Prediction and Filtering Theory", RIAS Technical Report 61-1, Baltimore, Maryland.

6. Lee, E.B., and Markus, L., "Optimal Control for Non Linear Processes", Arch. Rat. Mech. and Anal., 8,

7. Coddington, E.A., and Levinson, N.L., Theory of Ordinary Differential Equations, McGraw-Hill Book Co., Inc., New York, 1955.

8. Hestenes, M., Calculus of Variations and Optimal Control Theory, Wiley and Sons, 1966.

9. McShane, E.J., Integration, Princeton University Press, Princeton, N.J., 1944.

10. Bellman, R.E., "Upper and Lower Bounds for the Solution of the Matrix Riccati Equation", Journal of Math. Anal and Applic. 17 (1967), pp. 373-379.

11. Athans, M. and Schweppe, F.C., "On Optimum Waveform Design via Control Theoretic Concepts", Preprints J.A.C.C., Seattle, Washington, August, 1966, pp. 714-719.

REFERENCES

12. Rice, S.O., "Mathematical Analysis of Random Noise ", from the book Selected Papers on Noise and Stochastic Processes, Dover, Edited by Nelson Wax.

13. Wong, E., and Zakai, M., "On the Relation Between Ordinary and Stochastic Differential Equations", International Journal of Engineering Science, 3 (1965), pp. 213-229.

14. Lear, W.M., "SVEAD Users' Manual", TRW Systems Technical Report No. 7221.11-10, Redondo Beach, California.

15. Reid, W.T., "Some Remarks on Linear Differential Systems", Bulletin A.M.S., Vol. 45(1939), pp. 414-419.

16. Balakrishnan, A.V., "On a New Computing Technique in Optimal Control Theory and the Maximum Principle", Proc. of the Nat. Acad, of Sciences, January, 1968.

APPENDIX "A"

Discussion of the Correlated Noise Model

We consider stationary, correlated gaussian random processes, with band-pass power spectra of the form shown in Figure A.1. Such processes are typical of many physical systems - in particular the vibrations of airplane wings and rocket vehicles. In this appendix we shall be concerned with the selection of shaping filters for use in the Minimum Variance Estimator equations, the step-size for faithful numerical integration, and the digital generation of the random processes. This appendix is included because certain of the results are used in previous chapters, and it provides practical information which is not easily discerned from the literature.

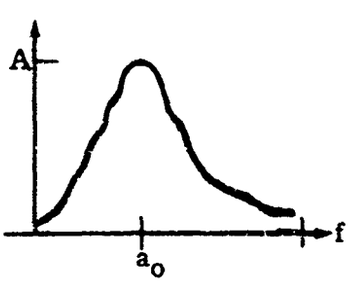


Figure A.1 Experimental Spectrum

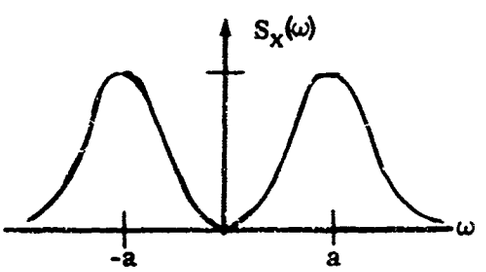


Figure A.2 Approximating Spectrum

A.1 Shaping Filter

In order to determine a suitable shaping filter, the experimental power spectrum shown in Figure A.1 needs to be approximated by an analytical function of the form shown in Figure A.2.

Power spectrums which might be used to represent the given power spectrum $S_x(\omega)$ are

$$S^{(1)}(\omega) = A^2 \left\{ \exp - \frac{(\omega+a)^2}{2\sigma^2} + \exp \frac{(\omega+a)^2}{2\sigma^2} \right\} \qquad (A.1)$$

and

$$S^{(2)}(\omega) = \frac{A^2}{(\omega-a)^2+b^2} + \frac{A^2}{(\omega+a)^2+b^2} \qquad (A.2)$$

We next approximate the given power spectrum (based on assumption and/or experimental data) by the fourth order rational function, and then synthesize a shaping filter for this rational spectrum using frequency domain techniques.

It should be remarked that the purpose of the shaping filters discussed here are not for generating random variables. They are used to give our problem the required canonic structure, and thus the approximations made here are not expected to be as critical as in the case of the shaping filters used for generating the correlated random process (Section A.2).

Given the experimental data, $S_{\tilde{x}}(\omega)$, the fourth order approximating spectrum, $S_x(\omega)$, would be obtained by minimizing (A.3) using standard approximating techniques,

$$\epsilon = \| S_{\tilde{x}} - S_x \|, \qquad (A.3)$$

where the norm $\|\cdot\|$ might be of the form

$$\int_0^{\omega_o} \left(S_{\tilde{X}}(\omega) - S_X(\omega)\right)^p dw \qquad 1 \leq p < \infty \tag{A.3a}$$

or

$$\sup_{0 \leq w \leq \omega_o} |S_{\tilde{X}}(\omega) - S_X(\omega)| \qquad p = \infty \tag{A.3b}$$

Of course the minimization is with respect to the parameters A, a, b, in Equation (A.2). We next demonstrate several properties of the assumed shaping filter.

Property 1 A shaping filter for the spectrum (A.2) is given by

$$H_X(s) = \frac{\sqrt{2}\,A\,(s + c)}{(s + b)^2 + a^2} \tag{A.4}$$

Demonstration

$$S_x(s) = \frac{2A^2 - s^2 + b^2 + a^2}{s^4 + 2s^2\,(a^2 - b^2) + (a^2 + b^2)^2}$$

Letting

$$\theta = \mathrm{Tan}^{-1}\,\frac{2ab}{-(a^2 - b^2)}, \quad \text{and} \quad c^2 = a^2 + b^2\ ,$$

there results

$$S_x(s) = \frac{2A^2\,(-s + c)(s + c)}{(s^2 + c^2\,e^{j\theta})\,(s^2 + c^2\,e^{-j\theta})}$$

whose pole-zero pattern is shown below

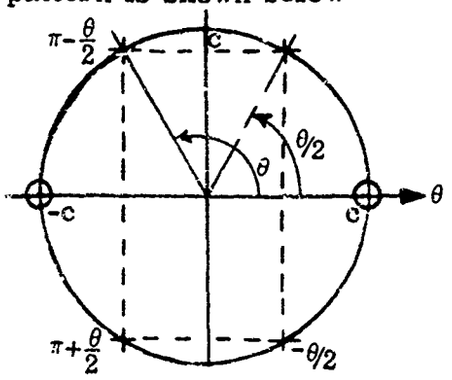


Figure A.3 Pole-Zero Pattern of $S_{\tilde{X}}(\omega)$

The transfer function of the shaping filter $H_X(s)$ is obtained from the left-half plane poles and zeros of $S_X(s)$, since we may write

$$S_X(s) = |H_X(s)|^2 \; S_W(s)$$

where $S_W(s) = 1$ is the power density spectrum of the white noise input to the filter, and $|H_X(s)|^2 = H_X(s)\; H_X(-s)$. Thus there results

$$H_X(s) = \frac{\sqrt{2}\,A\;(s + c)}{(s + ce^{j\theta/2})(s + ce^{-j\theta/2})}$$

and (A.4) follows by substituting for c and θ.

If Equation (A.4) is written in the time domain, we obtain

$$\ddot{x} + 2b\dot{x} + (a^2 + b^2)\; x = \sqrt{2}\,A\,\dot{w} + \sqrt{2}\,A\,cw$$

Thus, the representation of x as given above is not Markovian.

because it requires that the white noise w be differentiated. In the following property a method for rearranging the transfer function of Equation (A.4) is given so that a first-order Markovian process results.

Property 2 — By using the arrangement $H_{\hat{x}}(s)$ shown below in Figure A.4 a first order Markovian form results. Furthermore, this form may be written as

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -b & a \\ -a & -b \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \frac{A}{\sqrt{2}}\begin{bmatrix} 1 \\ \frac{c-b}{a} \end{bmatrix} w \qquad (A.5)$$

where $x \equiv 2y_1$, and A, a, b, c are as defined above.

Demonstration — From Figure A.4 we have

$$\hat{x} = x_1 + x_2$$

where $$\dot{x}_1 = -\frac{c_1}{b_1} x_1 + \frac{a_1}{b_1} w \qquad (A.5a)$$

and $$\dot{x}_2 = -\frac{c_2}{b_2} x_2 + \frac{a_2}{b_2} w$$

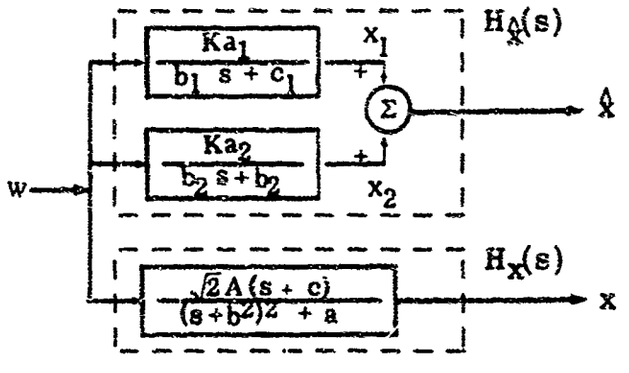


Figure A.4
Rearrangement of the Shaping Filter

The random variable $\hat{x}$ has the Markovian property which we are seeking, and the constants a_1, b_1, c_1, a_2, b_2, and c_2 are chosen such that x has the same power spectrum as x. Now

$$H_{\hat{x}}(s) = \frac{a_1}{b_1 s + c_1} + \frac{a_2}{b_2 s + c_2} =$$

$$= K\left(\frac{a_1 b_2 + a_2 b_1}{b_1 b_2}\right) \frac{s + \dfrac{a_1 c_2 + a_2 c_1}{a_1 b_2 + a_2 b_1}}{s^2 + \dfrac{c_1 b_2 + c_2 b_1}{b_1 b_2} s + \dfrac{c_1 c_2}{b_1 b_2}}$$

and equation (A.4) can be written as

$$H_x(w) = \frac{\sqrt{2}A\,(s + c)}{s^2 + 2bs + (a^2 + b^2)}$$

Equating coefficients in the equations for $H_{\hat{x}}$ and H_x gives

$$K\,\frac{a_1 b_2 + a_2 b_1}{b_1 b_2} = \sqrt{2}A, \quad \frac{a_1 c_2 + a_2 c_1}{a_1 b_2 \quad a_2 b_1} = c,$$

$$\frac{c_1 c_2}{b_1 b_2} = a^2 + b^2, \qquad \frac{c_1 b_2 + (c_2 b_1)}{b_1 b_2} = 2b.$$

From these relations there results

$$c_1/b_1 = b + ja \quad \text{and} \quad c_2/b_2 = b - ja$$

After further substitution there results

$$b_1 = \bar{b}_2, \quad c_1 = \bar{c}_2 = b_1(b + ja), \text{ and } b_1 = 1 + j\,\frac{b}{a}(1 \pm c/b) \tag{A.5b}$$

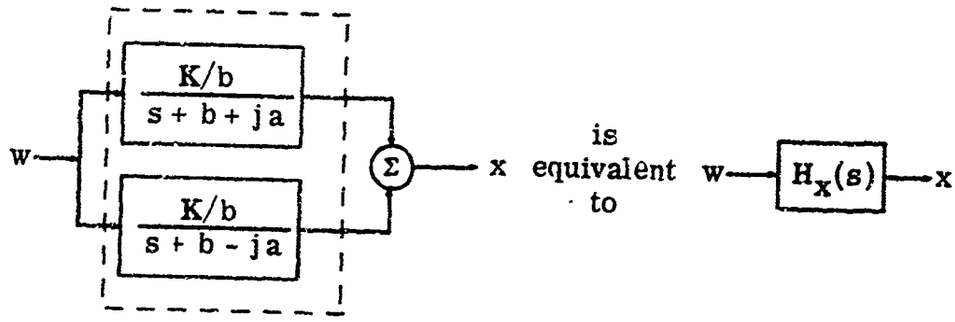
Equations (A.5b) imply that (choosing the negative sign)

$$H_x(s) = \frac{K}{b_1 s + c_1} + \frac{K}{b_2 s + c_2} = \frac{2K}{|b_1|^2} \frac{s + c}{s^2 + 2bs + c^2}$$

so that

$$K = \frac{A}{\sqrt{2}} \left[1 + \frac{1}{a^2}(b - c)^2\right]$$

This means that



Summarizing, we now have $x = x_1 + x_2$, where

$$\dot{x}_1 = -(b + ja)\, x_1 + \frac{K}{b_1} w$$

$$\dot{x}_2 = -(b - ja)\, x_2 + \frac{K}{b_1} w \qquad \text{(A.5c)}$$

b_1 and K are constants specified above.

From the symmetry in equations (A.5c) it is seen that $x_1 = x_2$ so that

$$\dot{x} = \dot{x}_1 + \dot{x}_2 = -(b+ja)\, x_1 - (b - ja)\, x_1 + K\left[\frac{1}{b_1} + \frac{1}{b_1}\right] w$$

$$= -2\mathrm{Re}\left[(b + ja)\, x_1\right] + K \frac{(b_1 + b_1)}{|b_1|^2} w,$$

and

$$\dot{y}_1 = -(y_1 b - a y_2) + \frac{A}{\sqrt{2}} w, \qquad \text{(A.5d)}$$

where $x_1 \triangleq y_1 + j y_2$. Equation (A.5c) implies that

$$\dot{y}_2 = -(a y_1 + b y_2) - \frac{A}{\sqrt{2}} \frac{b}{a} (1 - \frac{c}{b}) w \qquad \text{(A.5e)}$$

Equations (A.5d) and (A.5e) may be combined to give

$$\frac{d}{dt} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -b & a \\ -a & -b \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \frac{A}{\sqrt{2}} \begin{bmatrix} 1 \\ \frac{c-b}{a} \end{bmatrix} w \qquad \text{(A.5)}$$

where $x = \hat{x} = 2y_1$. This is the required Markov form for the random process x which was to be demonstrated.

Remark: The process specified by Equation (A.5) has the discrete representation

$$\begin{pmatrix} x/2 \\ y \end{pmatrix}_{k+1} = e^{-b\Delta} \begin{pmatrix} \cos a\Delta & \sin a\Delta \\ -\sin a\Delta & \cos a\Delta \end{pmatrix} \begin{pmatrix} x/2 \\ y \end{pmatrix}_k + \frac{A}{c} \sqrt{\frac{B}{2\Delta}} \begin{pmatrix} e^{-b\Delta} (\frac{c-b}{a} \sin a - \cos a) + 1 \\ e^{-b\Delta} (\frac{c-b}{a} \cos a + \sin a) + (\frac{b-c}{a}) \end{pmatrix} w_k \qquad \text{(A.6)}$$

where $\Delta = t_{k+1} - t_k$, $w_k = w(t)$ for $t_k \le t \le t_{k+1}$.

Property 3 The equation (A.7) below may also be used as a representation of the random process x.

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -c^2 & -2b \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \sqrt{2}A \begin{bmatrix} 1 \\ c-2b \end{bmatrix} w \qquad (A.7)$$

Demonstration Let

$$x_1 = x \qquad (A.7a)$$

$$\dot{x}_1 = x_2 + \sqrt{2}\, A\, w \qquad (A.7b)$$

$$\dot{x}_2 = -2b\, x_2 - (a^2 + b^2)\, x_1 + \sqrt{2}\, A\, (c-2b) w \qquad (A.7c)$$

Substituting (A.7a) into (A.7b) gives

$$x_2 = \dot{x} - \sqrt{2}\, A\, w$$

which upon substitution into equation (A.7c) gives

$$\ddot{x} - \sqrt{2}A\dot{w} = -2b\dot{x} + 2b\sqrt{2}\, A\, w - c^2 x + \sqrt{2}A(c-2b)\, w$$

$$= -2b\dot{x} - c^2 x + \sqrt{2}\, A\, c\, w \qquad (A.7d)$$

This equation has the vector form of Equation (A.7)

Remark This procedure is certainly shorter than the method given in Property 2, however, the substitutions are not obtained in a straightforward manner.

A.2 Digital Generation of a Correlated Random Process Specified by an Arbitrary Band-Pass Power Spectrum

A method for generating random variables digitally is to take linear combinations of outputs from some "standard" uncorrelated random number generator. In this Section we discuss a method for computing the coefficients in the linear combination of uncorrelated random numbers. The coefficients are chosen so that the resulting random variable has a power spectrum which approximates the given power spectrums in a mean square sense.

A white noise sequence sampled at unity time intervals has the spectral representation

$$w_j = \int_{-\pi}^{\pi} e^{i\omega j} \, dZ_w(\omega) \qquad (A.8)$$

where

$$E\left[dZ_w(\omega)\right] = 0 \text{ and } E\left|dZ_w(\omega)\right|^2 = d\omega,$$

$$dZ_w(-\omega) = \overline{dZ_w(\omega)},$$

$$E\, dZ_w(\omega)\ \overline{dZ_w(\omega')} = 0 \text{ if } \omega \neq \omega' \text{ (independant increments).}$$

In order to generate this correlated noise, we do not need a physically realizable scheme, i.e., we can use future values of

the random sequence, since these would be available in the computer memory.

Property 4 Suppose we assume that x_j can be approximated satisfactorily by a polynomial of the form

$$\tilde{x}_j = a_{-N} w_{j-N} + a_{-N+1} w_{j-N+1} +$$

$$+ \ldots + a_{-1} w_{j-1} + a_o w_j + a_1 w_{j+1} +$$

$$+ \ldots + a_{N-1} w_{j+N-1} + a_{j+N} w_{j+N} =$$

$$= \sum_{k=-N}^{+N} a_k w_{j-k}. \qquad (A.9)$$

It is assumed that $a_k = a_{-k}$ (corresponding future and past values are equally weighed), then the coefficients which approximate the given power spectrum in a least squares sense are given by

$$a_k = \frac{\Delta t}{\pi} \int_0^{\pi/\Delta t} \sqrt{S_x(\omega)} \cos k\Delta t\omega \, d\omega, \quad k=0, 1, \ldots, N. \qquad (A.10)$$

Demonstration Applying Equation (A.8) to (A.9) gives

$$x_j = \int_{-\pi}^{\pi} \left[\sum_{k=-N}^{N} a_k e^{i\omega k} \right] e^{i\omega j} dZ_w(\omega) = \qquad (A.10a)$$

$$= \int_{-\pi}^{\pi} \left[a_o + \sum_{k=1}^{N} 2a_k \cos k\omega \right] e^{i\omega j} dZ_w(\omega) = \int_{-\pi}^{\pi} e^{i\omega j} dZ_x(\omega)$$

Thus

$$E\left|dZ_x(\omega)\right|^2 = \left|a_o + 2\sum_{k=1}^{N} a_k \cos k\omega\right|^2 E\left|dZ_w(\omega)\right|$$

or

$$S_x(\omega)\, d\omega = \left|a_o + 2\sum_{k=1}^{N} a_k \cos k\omega\right|^2 d\omega,$$

and

$$S_x(\omega) = \left|a_o + 2\sum_{k=1}^{N} a_k \cos k\omega\right| \qquad \text{(A.10b)}$$

where $S_x(\omega)$ is the given power spectrum of the correlated noise x.

The quantity in the brackets of (A.10b) is real, so that we can write

$$S_x(\omega) = \left(a_o + 2\sum_{k=1}^{N} a_k \cos k\omega\right)^2 \qquad \text{(A.10c)}$$

We shall pick a mean-square error criterion for evaluating the coefficients a_k. However, in order to obtain simplified expressions for these coefficients, we approximate $\sqrt{S_x(\omega)}$ instead of $S_x(\omega)$. We then obtain the mean-square error over the interval $[-\pi, \pi]$ as

$$\epsilon = \int_{-\pi}^{\pi}\left[\sqrt{S_x(\omega)} - 2\sum_{k=0}^{N}{}' a_k \cos k\omega\right]^2 d\omega$$

(Note: The prime on the summation indicates a_o is divided by two).

$$= \int_{-\pi}^{\pi} S_x(\omega)\, d\omega + 4\sum_{k=0}^{N}{}' a_k^2 \int_{-\pi}^{\pi} \cos^2 k\omega\, d\omega -$$

$$-4\sum_{k=0}^{N}{}' a_k \int_{-\pi}^{\pi} \sqrt{S_x(\omega)}\cos k\omega\, d\omega + 4\sum_{\substack{k=0 \\ k\neq j}}^{N}\sum_{j=0}^{N} a_k a_j \int_{-\pi}^{\pi} \cos k\omega \cos j\omega\, d\omega$$

Now in actual fact, the power spectrum $S_x(\omega)$ will be zero after some sufficiently large angular frequency $\omega = \Omega$, and the sampling frequency would not be unity, but $1/\Delta t$, where Δt is the time (in seconds) between samples. Thus, instead of $\cos k\omega$, we have $\cos k\Delta t\omega$. If we choose the interval $[-\pi/\Delta t,\ \pi/\Delta t]$ and the sampling period Δt such that $\pi/\Delta t > \Omega$ and Δt is compatible with integration requirements (next section) then

$$\int_{-\pi/\Delta t}^{\pi/\Delta t} \cos k\Delta t\omega \cos j\Delta t\omega\, d\omega = \begin{cases} \frac{\pi}{\Delta t} & j = k \\ 0 & j \neq k \end{cases}$$

The error ϵ is then

$$\epsilon = \int_{-\Omega}^{\Omega} S_x(\omega)\, d\omega + \frac{4\pi}{\Delta t}\sum_{k=0}^{N}{}' a_k^2 - 4\sum_{k=0}^{N}{}' a_k \int_{-\pi/\Delta t}^{\pi/\Delta t} \sqrt{S_x(\omega)}\cos k\Delta t\omega\, d\omega$$

Since we are minimizing ϵ with respect to the a_k, which are differentiable and unconstrained, we obtain

$$\frac{8\pi}{\Delta t}\ a_k - 4\int_{-\pi/\Delta t}^{\pi/\Delta t} \sqrt{S_x(\omega)}\cos k\Delta t\omega\, d\omega = 0 \quad k = 0, 1, \ldots, N.$$

or

$$a_k = \frac{\Delta t}{2\pi}\int_{0}^{\pi/\Delta t} \sqrt{S_x(\omega)}\cos k\Delta t\omega\, d\omega, \quad k = 0, 1, \ldots, N \qquad \text{(A.10d)}$$

We are sure that this is a minimum since the second derivatives with respect to a_k are > 0. Since the integrand in equation (A.10d) is an even function, we obtain (A.10) as required.

Property 5 The percent error in the representation specified by Property 4 is given by

$$\left| 1 - \frac{4\pi}{P\Delta t} \left[\frac{a_0^2}{2} + \sum_{k=1}^{N} a_k^2 \right] \right| \qquad (A.11)$$

where P is the power in the given random process.

Demonstration Since $\sqrt{S_x(\omega)}$ is an even function, it has a Fourier Series expansion

$$\sqrt{S_x(\omega)} = \frac{\alpha_0}{2} + \sum_{k=1}^{\infty} \alpha_k \cos \frac{k\pi}{\Gamma} \omega, \quad \omega \epsilon \left[-\Gamma, \Gamma \right], \qquad (A.11a)$$

where

$$\alpha_k = \frac{2}{\Gamma} \int_0^{\Gamma} \sqrt{S_x(\omega)} \cos \frac{k\pi}{\Gamma} \omega \, d\omega, \quad k = 0, 1, \ldots \qquad (A.11b)$$

If we compare equations (A.11b) and (A.10) with $\pi/\Delta t = \Gamma$, we see that $a_k = \alpha_{k/2}$. Thus $\sqrt{S_x(\omega)} = a_0 + 2 \sum_{k=1}^{\infty} a_k \cos k\Delta t \omega$, and the power in the random process is

$$P = 2\int_0^{\pi/\Delta t} S_x(\omega)\, d\omega = 2\int_0^{\Omega} S_x(\omega)\, d\omega = \int_{\pi/\Delta t}^{\pi/\Delta t} (a_0 + 2\sum_{k=1}^{\infty} a_k \cos k\Delta t\omega)^2 d\omega =$$

$$= \int_{\pi/\Delta t}^{\pi/\Delta t} a_0^2 + 4\sum_{k=1}^{\infty} a_k^2 \int_{-\pi/\Delta t}^{\pi/\Delta t} \cos^2 k\Delta t\omega \, d\omega = \frac{4\pi}{\Delta t} \left[\frac{a_0^2}{2} + \sum_{k=1}^{\infty} a_k^2 \right] \qquad (A.11c)$$

Letting

$$\tilde{P} = \frac{4\pi}{\Delta t}\left[\frac{a_0^2}{2} + \sum_{k=1}^{N} a_k^2\right] \qquad \text{(A.11d)}$$

the error in total power between the given random variable x, and the synthesized random variable $\tilde{x}$ is

$$\Delta P = |P - \tilde{P}| = \left|2\int_0^{\Omega} S_x(\omega)\, d\omega - \frac{4\pi}{\Delta t}\left[\frac{a_0^2}{2} + \sum_{k=1}^{N} a_k^2\right]\right| \qquad \text{(A.11e)}$$

This error is a function of the number of coefficients used, N, and the error is inversely proportional to N. This monotonicity of the truncated Fourier series approximation is due to the orthogonality of the cosine functions. As a percentage, Equation (A.11e) becomes

$$\Delta P/P = \left|1 - \frac{4\pi}{P\Delta t}\left[\frac{a_0^2}{2} + \sum_{k=1}^{N} a_k^2\right]\right| \qquad \text{(A.11f)}$$

the required result (A.11).

Remark: Suppose it is required that the coefficients a_k and N are chosen so that the total power of $\tilde{x}$ be within α of the power in x, that is,

$$\left|\frac{P - \tilde{P}}{P}\right| \leq \alpha \qquad \text{(A.11g)}$$

This could be a good criterion for determining the number of coefficients N.

Computational Results

The two power spectrums specified below by Equations (A.12a) and (A.12b) were considered.

$$S^{(1)}(\omega) = \frac{A^2}{\sigma\sqrt{2\pi}} \left\{ \exp \frac{-(\omega - a)^2}{2\sigma^2} + \exp \frac{-(\omega + a)^2}{2\sigma^2} \right\} \qquad \text{(A.12a)}$$

and

$$S^{(2)}(\omega) = \frac{A^2}{(\omega - a)^2 + b^2} + \frac{A^2}{(\omega + a)^2 + b^2} \qquad \text{(A.13a)}$$

The corresponding autocorrelation functions are, respectively,

$$R^{(1)}(\tau) = 2A^2 e^{-\tau^2 \sigma^2/2} \cos a\tau \qquad \text{(A.12b)}$$

and

$$R^{(2)}(\tau) = \frac{A^2}{b} e^{-b|\tau|} \cos a\tau \qquad \text{(A.13b)}$$

Computational results for $S^{(1)}(\omega)$ with a = 24.0 radians, σ = 10.0, and $S^{(1)}(a) = 1.0$, and for $S^{(2)}(\omega)$, with a = 24.0 radians, b = 5.0, and $S^{(2)}(a) = 1.0$, are shown below. In these tables

$$\alpha = \left| 1 - \frac{4\pi}{P\Delta t} \left[\frac{a_0^2}{2} + \sum_{k=1}^{N} a_k^2 \right] \right| .$$

TABLE 1 SAMPLING TIME $\Delta t = \pi/\Omega = \pi/75 = .0427$ sec

	$S^{(1)}(\omega)$		$S^{(2)}(\omega)$	
N	a_k	α	a_k	α
0	.455257*		.387541*	
1	.196292	.14930	.113449	.18522
2	-.113420	.07231	-.066344	.14067
3	-.106863	.00398	-.095744	.04768
4	-.025387	.00012	-.045514	.02683
5	-.002984	.00007	-.016735	.02399
6	-.002264	.00004	.037336	.00988
7	-.001855	.00002	.021766	.00513
8	-.001288	.00001	-.004489	.00487

*This coefficient has <u>not</u> been divided by two.

The two sets of curves on the following pages show $S^{(1)}(\omega)$ and $S^{(2)}(\omega)$ and the approximating power spectrums

$$S_x^{(1)}(\omega) = (a_0 + 2 \sum_{k=1}^{N} a_k \cos k\Delta t\omega \quad , \quad i = 1, 2,$$

for N = 10 and 15, and $\Delta t = 0.02$ seconds (That is, the curves corresponding to the coefficients of Table 2).

From these curves we see that the exponential power spectrum, $S^{(1)}(\omega)$, is reasonably well approximated as a function of ω by N = 10, and very well by N = 15. On the other hand, the rational power spectrum, $S^{(2)}(\omega)$ is not well approximated, even by N = 15. Thus N depends on t and on the shape of the given power spectrum.

TABLE 2 SAMPLING TIME $\Delta t = 0.02$ sec

	$S^{(1)}(\omega)$		$S^{(2)}(\omega)$	
N	a_k	α	a_k	α
0	.217369*		.175520*	
1	.184170	.27880	.139382	.26105
2	.102510	.14710	.061514	.18285
3	.013242	.14490	-.002265	.18066
4	-.047379	.11677	-.029463	.16222
5	-.067106	.06033	-.037714	.13207
6	-.056730	.01999	-.044446	.09016
7	-.035391	.00430	-.044527	.04810
8	-.170496	.00065	-.029811	.02924
9	-.006412	.00014	-.008786	.02761
10	-.002095	.00008	.004843	.02711
11	-.001062	.00007	.010375	.02482
12	-.001053		.015116	.01998
13	-.001071		.018987	.01233
14	-.000968		.016050	.00686
15	- 000845		.007177	.00577

*Coefficient has not been divided by two.

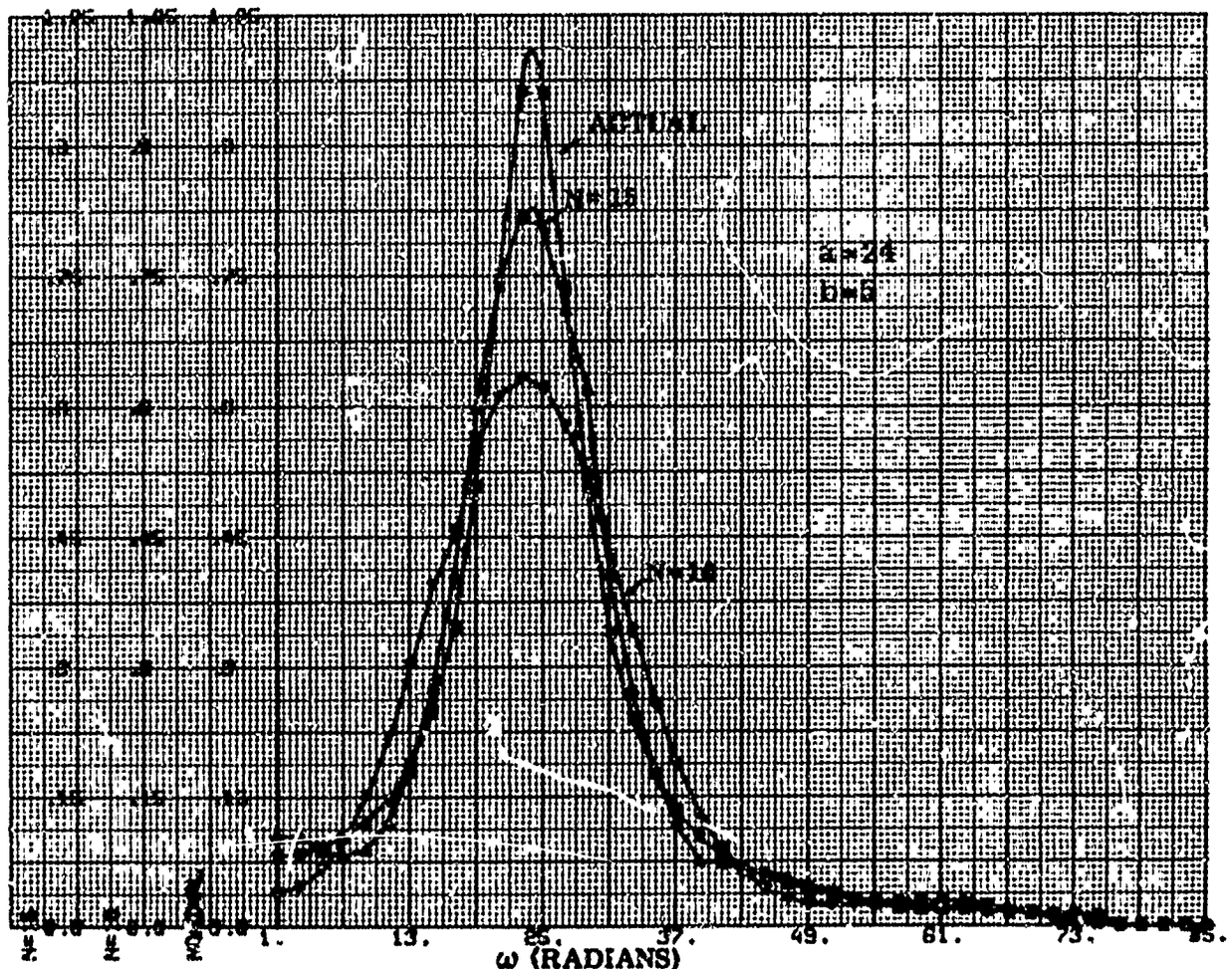


Figure A.5a Approximation of Rational Spectrum

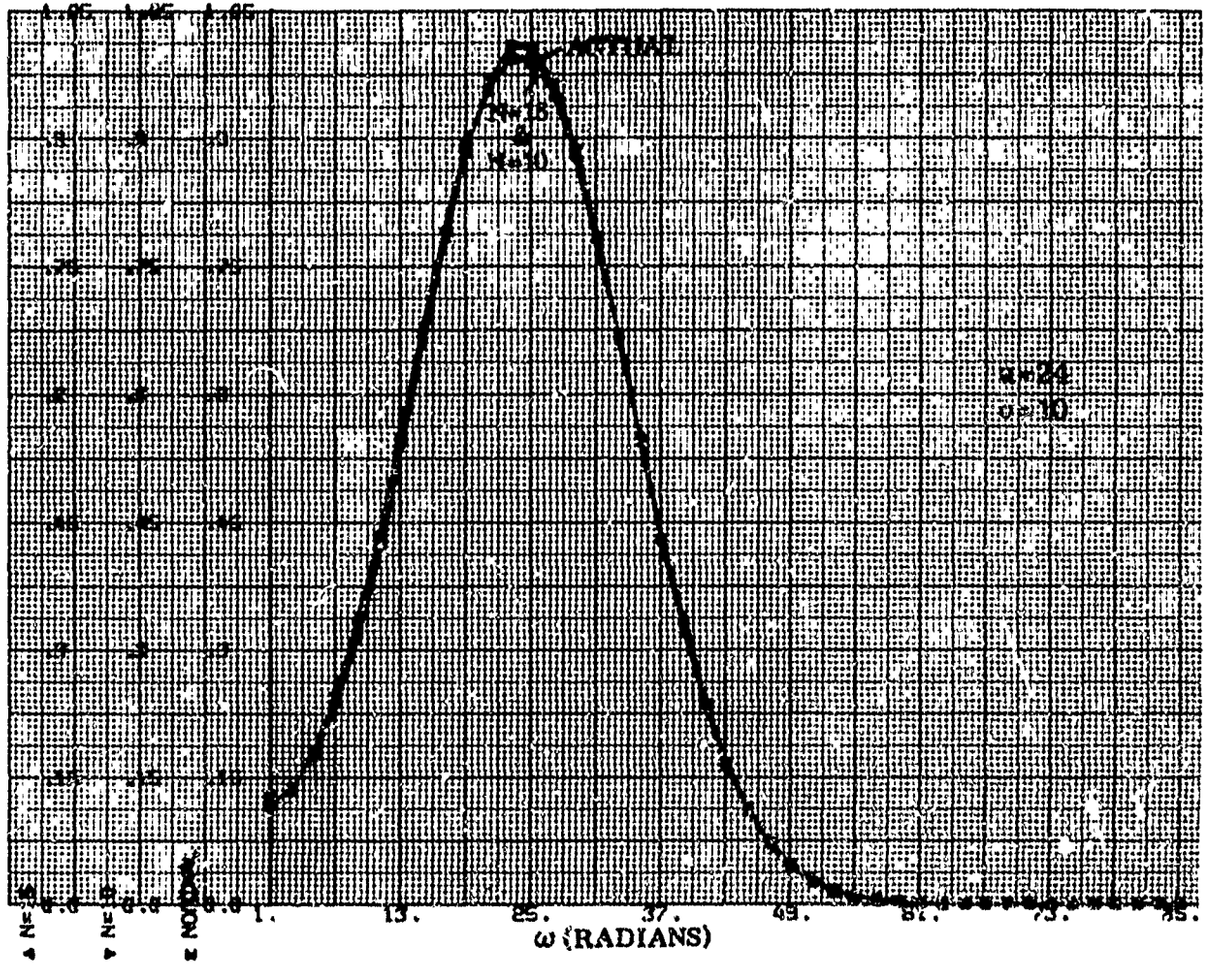


Figure A.5b Approximation of Exponential Spectrum

A.3 Step Size for Digital Generation and Integration of the Correlated Random Process

We consider a zero-mean stationary, gaussian random process with power spectrum $S(\omega)$ and autocovariance function $R(\tau)$ given by

$$R(\tau) = \frac{1}{2\pi}\int_0^\infty S(\omega)\cos\omega\tau\, d\omega$$

where it is noted that here the one-sided spectrum $S(\omega)$ is used.

A good indication of the step sizes can be obtained if the number of zero-crossings per unit time the number of maximum and minimum per unit time and the r.m.s. value of the derivative are known. For stationary, gaussian, random processes the expected values of these quantities can be calculated relatively easily in terms of the power spectrums or autocovariance functions. Since the random process is assumed zero-mean and gaussian, the probability distribution is completely specified by $R(0)^{1/2}$. That is, the probability that $x(t)$ lies between x and $x+dx$ is

$$\frac{dx}{\sqrt{2\pi R(0)}}\ \exp - x^2/2R(0).$$

Thus the probability distributions of such processes are completely specified by their variance.

The expected number of zeros per second of x(t) is given by (Equation(3.3-11) of [12])

$$\lambda_2 \equiv E\left[\text{No. of zeros/sec}\right] = \frac{1}{\pi}\left[\frac{-R''(0)}{R(0)}\right]^{1/2} = \frac{1}{\pi}\left[\frac{\int_0^\infty \omega^2 S(\omega)d\omega}{\int_0^\infty S(\omega)d\omega}\right]^{1/2} \tag{A.14}$$

The expected number of maxima and minima is given by (Equation (3.6-6 of

$$\lambda_4 = E\left[\text{No. of max and min per sec}\right] = \frac{1}{\pi}\left[\frac{-R^{(iv)}(0)}{R''(0)}\right]^{1/2} =$$

$$= \frac{1}{\pi}\left[\frac{\int_0^\infty \omega^4 S(\omega)\,d\omega}{\int_0^\infty \omega^2 S(\omega)\,d\omega}\right]^{1/2} \tag{A.15}$$

The r.m.s. value of x(t) is given by

$$\sigma_{\dot{x}} = \left[R_{\dot{x}}(0)\right]^{1/2} = \left[\frac{1}{2\pi}\int_0^\infty \omega^2 S(\omega)\,d\omega\right]^{1/2} \tag{A.16}$$

where

$$R^{(n)}(0) = \frac{d^n}{d\tau^n} R(\tau)\Big|_{\tau=0} \tag{A.17}$$

Remark In the above expressions it is assumed that the necessary derivatives of $R(\tau)$ in the neighborhood of $\tau=0$ exist. According to Theorems VIII and XI in section 20.5 of A.E. Taylor's Advanced

Calculus, the necessary derivatives of $R(\tau)$ exist in the neighborhood of $\tau=0$ provided that the integrals in equations (A.14), (A.15), (A.16) converge. This is a useful result as we do not need to have $R(\tau)$ explicitly. For example, in equations (A.13a) and (A.13b) above, $R_2(\tau)$ does not have the necessary derivatives at $\tau=0$. If we truncate $S_2(\omega)$ all of the necessary integrals in equations (A.14), (A.15) and (A.16) converge, thus we are sure by this result that the necessary derivaties of $R_2(\tau)$ corresponding to this truncated $S_2(\omega)$ exist. It would be fairly tedious to obtain the $R_2(\tau)$ corresponding to this truncated $S_2(\omega)$, and then to verify that the necessary derivatives exist.

Generally, the relations (A.14), (A.15) and (A.16) cannot be calculated readily. However, for the exponential power spectrum (A.12a) we may calculate the quantities given by (A.14), (A.15) and (A.16) in terms of the derivatives of $R_1(\tau)$ evaluated at $\tau=0$. In particular we have

Property 6 For the exponential spectrum, the asymptotic values ($\Omega \to \infty$) may be calculated from

$$E\left[\text{No. of zeros/sec}\right] = \frac{1}{\pi}\sqrt{\frac{-R''(0)}{R(0)}} = \frac{1}{\pi}\sqrt{a^2+\sigma^2} \qquad \text{(A.18a)}$$

$$E\left[\text{No. of max and min/sec}\right] = \frac{1}{\pi}\sqrt{\frac{a^4+6a^2\sigma^2+3\sigma^4}{a^2+\sigma^2}} \qquad \text{(A.18b)}$$

and

$$\sigma_{\dot{x}} = \left[R_{\dot{x}}(0)\right]^{\frac{1}{2}} = \left[-R''(0)\right]^{\frac{1}{2}} = \sqrt{2A(a^2+\sigma^2)} \qquad \text{(A.18c)}$$

Demonstration Letting

$$R(\tau)/2A = e^{-\tau^2\sigma^2/2}\cos a\tau \equiv f_1(\tau)\cdot f_2(\tau)$$

there results

$$R(0) = f_1(0)\,f_2(0) = 1$$

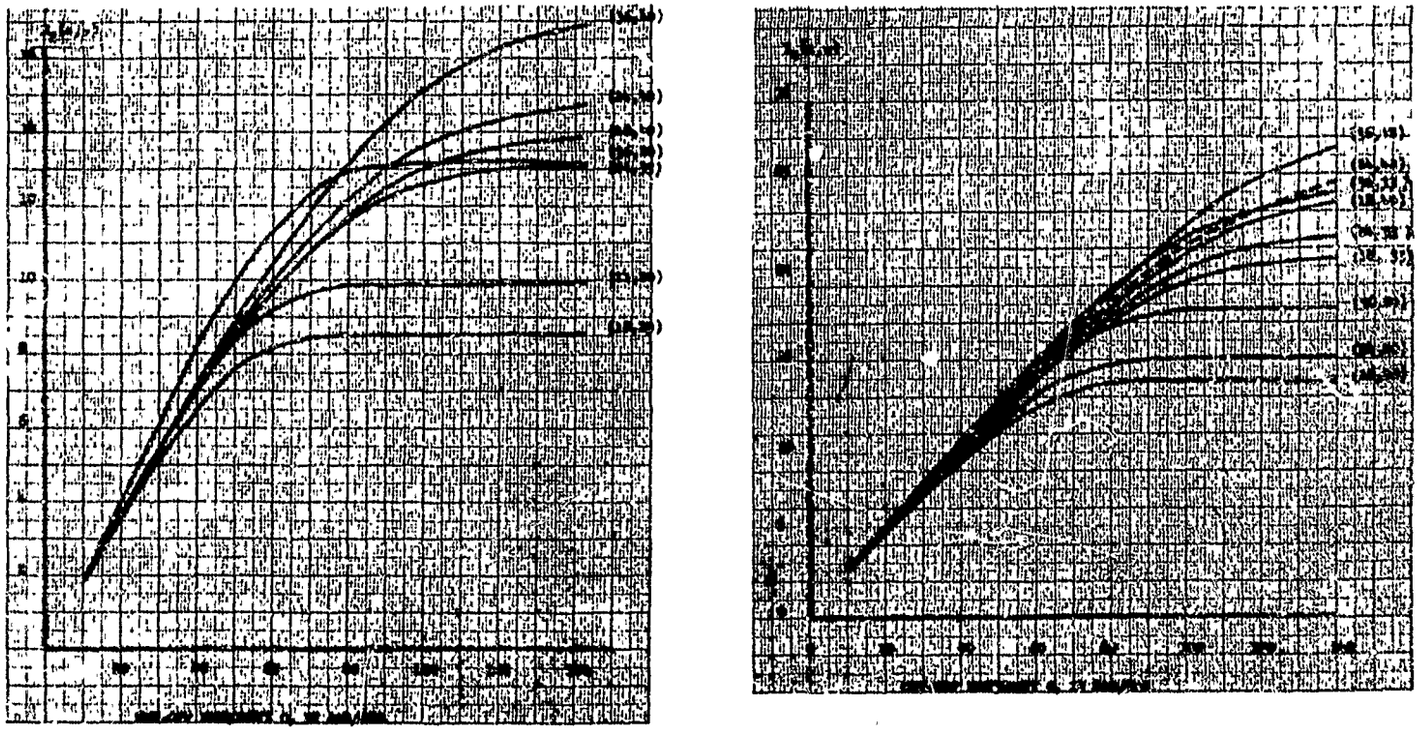
$$R''(0) = f_1(0)\,f_2(0)^{(')} + 2f_1(0)^{(')}f_2(0)^{(')} + f_1(0)^{(')}f_2(0) = -\sigma^2 - a^2$$

$$R^{(4)}(0) = f_1(0)f_2(0)^{(4)} + 4f_1(0)^{(1)}f_2(0)^{(3)} + 6f_1(0)^{(2)}f_2(0)^{(2)} + 4f_1(0)^{(3)}f_2(0)^{(1)} + f_1(0)^{(4)}f_2(0) = a^4 + 6a^2\sigma^2 + 3\sigma^4$$

Substituting these results into (A.14), (A.15) and (A.16) yields equations (A.18a), (A.18b), (A.18c), as required.

Experimental Results

Values of expressions (A.14), (A.15), and (A.16) for the two power spectrums (A.12a) and (A.13a), and for various values of the constants a, b, σ, and Ω are shown in the curves below. The asymptotic values for the exponential spectrum (Figure A.6) agree exactly with the values calculated from Equations (A.18a) and (A.18b) (that is, when the asymptotic values are reached in Figure A.6.



(a) Expected Number of Zero Crossings. (b) Expected Number of Maxima and Minima.

Figure A.6 Expected Number of Zero Crossings, I_2, and the Expected Number of Maxima and Minima, I_4, for the Exponential Spectrum.

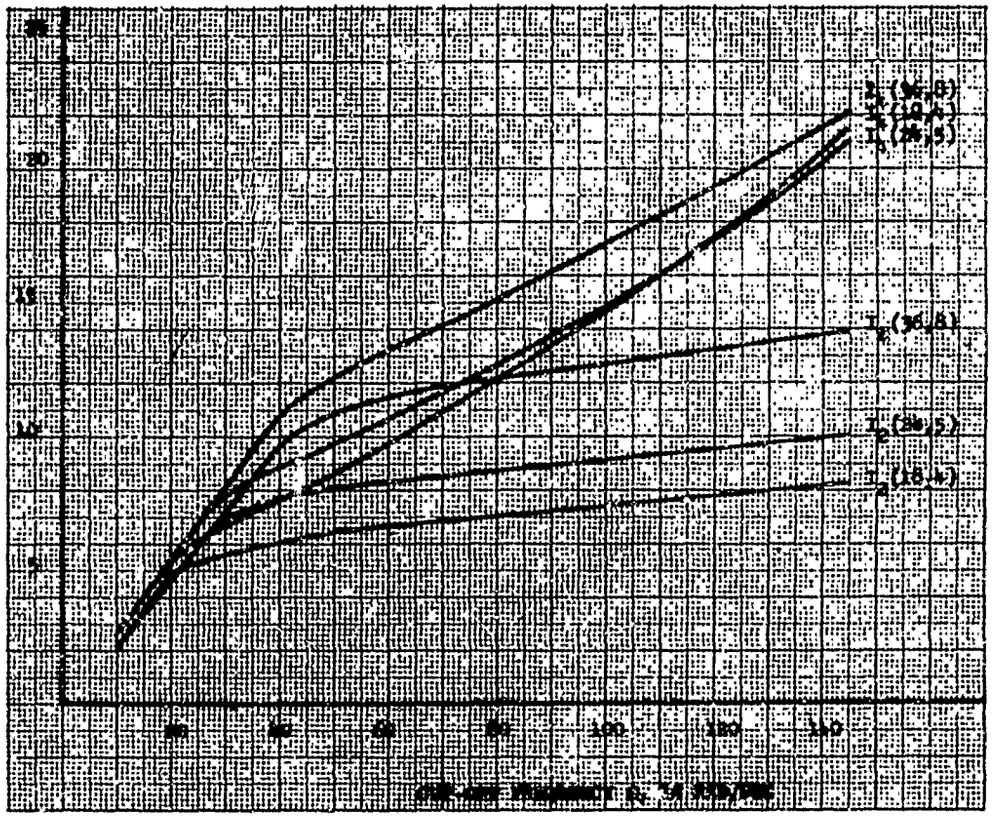
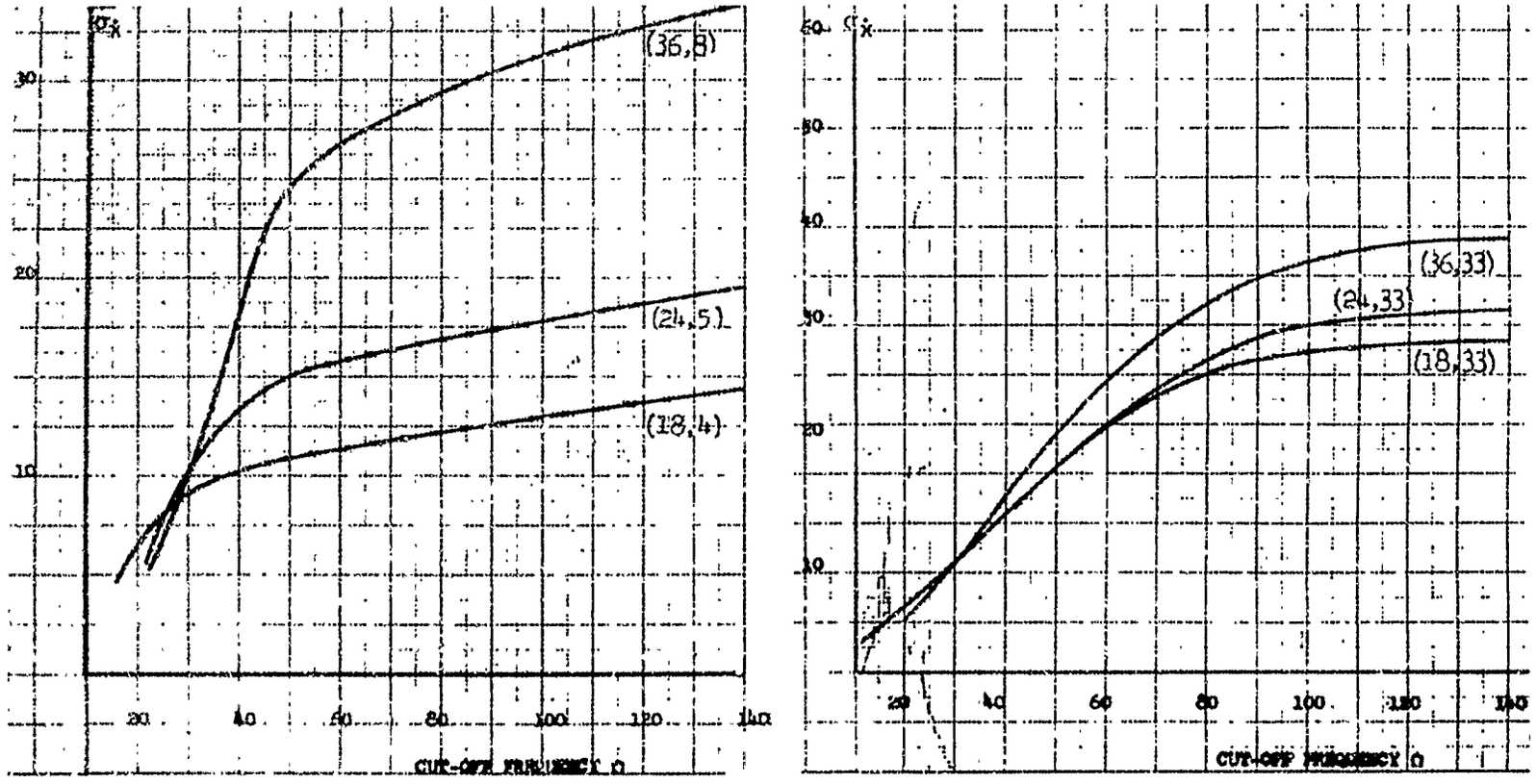


Figure A.7 Expected Number of Zero Crossings, I_2, and the Expected Number of Maxima and Minima, I_4, for the Rational Spectrum.



(a) Rational Power Spectrum (b) Exponential Power Spectrum
(NOTE: $S(\omega)$ is normalized so that $S(a) = 1.0$)

Figure A.8 Standard Deviation of the Random Process $\dot{x}$, $\sigma_{\dot{x}}$.

To give some physical feeling for the above discussion and results, a white noise sequence having a flat power spectrum to 100 cycles/second was generated (white noise was sampled 100 times/second). This sequence was then passed through the low-pass filter (0 to 70 radians/second). By actually counting the number of zero crossings and the number of maximum and minimum on the time histories the experimental results are compared below with those obtained from the integral equations (A.14) and (A.15).

Time-Interval	No. of Zeros	No. of Max & Min
0-1	11	18
1-2	14	18
2-3	14	20
3-4	17	17
4-5	14	19
5-6	15	22
6-7	13	14
7-8	13	17
8-9	14	17
9-10	13	17
TOTAL	138	179

These results show good correspondence between values calculated from Equations (A.14) and (A.15). In particular, the

average number of experimental zero crossings is 13.8, and the expected number of zero crossings from Equation (A.14) is 14.0. The average number of experimental maxima and minima is 17.9, and the expected number of maxima and minima from Equation (A.15) is 18.4.

It is important to note that the observed values of zero-crossings and maxima and minima do not vary drastically from the average values. This may be readily observed by comparing the columns of values above with the average values, 13.8 and 17.9, respectively.

A.4 Comparison of Three Methods for Generating Digitally Stationary Correlated Random Processes

In this section we compare the two models for generating the correlated random processes specified above by Equations (A.6) and (A.9), along with a third set of equations to be specified below. It is instructive to make such a comparison because each of the three shaping filters considered here has been formulated using different assumptions and approximations. Furthermore, each of the shaping filters would have advantages over the other two in the application of the minimum variance techniques.

The first filter, Filter I (Equation A.9)), is useful because arbitrary power spectra may be approximated without altering the filter design. Essentially this filter uses a moving average of uncorrelated random numbers to generate correlated noise. This method has the disadvantages that a digital subroutine is required to generate its coefficients, and that the filter is not in canonic form for use in the error-analysis equation of the minimum-variance estimator. However, it is useful for generating observations.

The second and third filters, Filter II (Equation (A.6)) and Filter III, both require that the given power spectra be first approximated

by a fourth order rational spectra. These filters use feedback to generate the desired correlated noise in terms of uncorrelated random numbers. Filters II and III generally would not represent the given power spectra as faithfully as Filter I. However, these filters are in the canonic form for use in the error-analysis equation. The third filter probably gives a more faithful representation of the specific rational spectrum considered than does Filter II, since the coefficients of Filter III have been adjusted to improve the approximation in going from a continuous random process to a discrete random process.

Filter II is used for the same purpose as Filter III and the derivation of Filter II is obtained in a straightforward fashion. The derivation of Filters I and II is given above, and the specification of Filter III is given in [14] . Combinations of the three above-mentioned shaping filters could be useful in studying the sensitivity of the minimum-variance estimates of the state vector due to variations in the state noise power spectra.

We give only a brief description of the three shaping filters considered. Additional details of these filters may be found above.

Filter I (Section A.2)

We are given a stationary, gaussian random process, x, with a power spectrum $S_x(\omega)$. We assume that $x(j\Delta t) \triangleq x_j$ may be approximated by

$$\tilde{x}_j = \langle \underline{a}, \underline{w}_j \rangle = \sum_{k=-N}^{N} a_k w_{j-k} \qquad (A.19)$$

where w_j is a white noise sequence and is obtained from a "standard" uncorrelated random number generator, and the coefficients a_k are constants which can be computed from the power spectrum $S_x(\omega)$ of x_j. Specifically, the a_k are given by Equation (A.10) and N is chosen to satisfy (A.11g).

Filter II (Section A.1)

We make the assumption that the given power spectrum may be suitably approximated by the fourth order rational power spectrum

$$S_x(\omega) = \frac{A^2 B}{(\omega - a)^2 + b^2} + \frac{A^2 B}{(\omega + a)^2 + b^2} \qquad (A.20a)$$

where B is chosen so that $S_x(\pm a) = A^2$. A model for this process is specified by Equation (A.5). If it is assumed that the noise w(t) is constant over the relatively short computation interval Δ, (in Section A.3 quantitative procedures are given

which insure that this assumption is realistic; the distinction between a white noise process and a white noise sequence will be discussed below), then the solution of (A.20b) may be written as*

$$\begin{pmatrix} \frac{x_i}{2} \\ y_i \end{pmatrix} = e^{-b\Delta} \begin{pmatrix} \cos a\Delta & \sin a\Delta \\ -\sin a\Delta & \cos a\Delta \end{pmatrix} \begin{pmatrix} \frac{x_{i-1}}{2} \\ y_{i-1} \end{pmatrix} + \\ + \frac{A}{c}\sqrt{\frac{B}{2\Delta}} \begin{pmatrix} e^{-b\Delta}\left(\frac{c-b}{a}\sin a\Delta - \cos a\Delta\right) + 1 \\ e^{-b\Delta}\left(\frac{c-b}{a}\cos a\Delta + \sin a\Delta\right) + \left(\frac{b-c}{a}\right) \end{pmatrix} w_{i-1} \qquad \text{(A.20b)}$$

Simplifying the notation, (A.20) may be written as

$$\underline{x}_i = A_{II}\,\underline{x}_{i-1} + \underline{b}_{II}\,w_{i-1} \qquad \text{(A.20)}$$

where the definition of $\underline{x}_i$, A_{II}, and $\underline{b}_{II}$ is obvious from (A.20). The "white noise" sequence w_{i-1} has the property

$$E\left[w_i\right] = 0 \text{ and } \mathrm{Cov}(w_i, w_j) = E\left[w_i\, w_j\right] = \delta_{ij} ,$$

where

$$\delta_{ij} = \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j \end{cases}$$

*Wong and Zakai, [13] , have shown that one generally can't replace a stochastic differential equation with the limit of solutions of an ordinary differential equation, as the solutions of the latter usually do not converge to the solution of the former. However, the solution (A.20b) would converge to that of (A.6) due to linearity of the equation.

Filter III

A discrete model of the process defined by (A.20a) is given in [14] as

$$x_i = b_1 x_{i-1} - b_2 x_{i-2} + a_1 w_{i-1} + a_2 w_{i-2} \qquad \text{(A.21a)}$$

where w_i is as above, and where

$$b_1 = 2e^{-b\Delta}\cos a\Delta \qquad b_2 = e^{-2b\Delta}$$

$$a_1 = \tfrac{1}{2}\sqrt{(1-b_2)\varphi_o}\left[\sqrt{1+b_2-b_1} - \sqrt{1+b_2+b_1}\right]$$

$$a_2 = \tfrac{1}{2}\sqrt{(1-b_2)\varphi_o}\left[\sqrt{1+b_2-b_1} + \sqrt{1+b_2+b_1}\right] \qquad \text{(A.21b)}$$

Δ = sampling interval $\varphi_o = \dfrac{A^2B}{b}$

Next let, $y_i \triangleq x_{i-1} - (a_2/b_2)\, w_{i-1}$ so that $x_i = b_1 x_{i-1} - b_2 y_{i-1} + a_1 w_{i-1}$. These equations may be written in matrix form as

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} b_1 & -b_2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix} + \begin{pmatrix} a_1 \\ \dfrac{-a_2}{b_2} \end{pmatrix} w_{i-1} \qquad \text{(A.21c)}$$

Again, to simplify the notation, we write (A.21c) as

$$\underline{x}_i = A_{III}\,\underline{x}_{i-1} + \underline{b}_{III}\, w_{i-1} \qquad \text{(A.21)}$$

Remark If a white noise process, $w(t)$ has covariance $R(t)\,\delta(t-\tau)$, then the corresponding white noise sequence is

$$w_j = w(t = j\Delta)/\sqrt{\Delta} \quad \text{and} \quad \text{Cov}\,(w_j) = R(t = j\Delta)/\Delta$$

The covariances for Filter I and II were assumed to be for a white noise process, and that for Filter III was for a white noise sequence. Thus suitable normalizing factors must be introduced into each.

Specific Parameters Used for the Numerical Comparison

In order to make a quantitive comparison of the three shaping filters discussed above, we took a = 24, b = 5 and A = 5. For these values of a and b, a sampling rate of Δ = .02 seconds was selected. This value of Δ was chosen according to the criteria given in Section A.3. For these values of a, b, A, and Δ, the coefficients in Equation(A.19) are given in Table II, Section A.2, and the matrices in Equations (A.20) and (A.21) become

$$A_{II} = \begin{pmatrix} .802586 & .417835 \\ -.417835 & .802586 \end{pmatrix} \qquad b_{II} = \begin{pmatrix} 2.724374 \\ 1.305003 \end{pmatrix}$$

$$A_{III} = \begin{pmatrix} 1.605172 & -.818731 \\ 1.0 & 0 \end{pmatrix} \qquad b_{III} = \begin{pmatrix} -3.286653 \\ -6.686670 \end{pmatrix}$$

Simulation Results

By actually counting the number of zero crossings and the number of maxima and minima on the time histories of simulated filters

the results of the three filters can be compared with those established in Section A.3. The observed values for a 5 second sample are shown below

	Numerical Results (a=24, b=5, Δ=.02)	
	Number of Zeros in 5 Seconds	Number of Maxima & Minima in 5 Seconds
White Noise	129	170
Filter I	50	111
Filter II	51	87
Filter III	47	96

From Section A.3, the expected number of zeros in 5 seconds is approximately 51, and the expected number of maxima and minima in 5 seconds is approximately 115 when sampled 50 times/second. For white noise sampled every .02 seconds, these numbers are 28.9/second and 33.6/second, respectively.

Since we are assuming gaussian random variables, it is of interest to count the number of points outside of ± 1 sigma:

	White Noise	Filter I	Filter II	Filter III
Percent of Total points outside of ± 1 sigma	33	24	32	28

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)	2a. REPORT SECURITY CLASSIFICATION
University of California School of Engineering and Applied Science Los Angeles, California 90024	UNCLASSIFIED 2b. GROUP

3. REPORT TITLE

ON THE IDENTIFICATION OF INERTIAL MEASUREMENT UNIT ERROR PARAMETERS

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific Interim

5. AUTHOR(S) (First name, middle initial, last name)

John Baziw

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO OF REFS
April 1970	113	0

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
AF-AFOSR 699-67 b. PROJECT NO. 9749-01 c. 61102F 681304 d.	70-32 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) AFOSR 70-1851 TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH, OTHER	Air Force Office of Scientific Research(SRM) 1400 Wilson Boulevard Arlington, Virginia 22209

13. ABSTRACT

The problem of identifying the error parameters associated with inertial measurement units is considered in this report. This is an important practical problem which is included in a large class of system parameter identification problems. A general approach for formulating the many possible inertial measurement unit (IMU) error parameter configurations is given, and specific realizations are specified in detail. The formulation is such that time correlated environmental and observational random disturbances can be incorporated. Experimental results showing the effects of state and observation noise power levels, and the nominal trajectory on the identification of the error parameters for three specific configurations are presented. These results indicate that a meaningful optimization problem can be formulated in terms of the nominal trajectory variables. The problem is then considered as an optimal control problem with the cost being a functional of the estimation covariance matrix and the controls, where certain nominal trajectory variables are considered as the controls. The question of the existence of optimal controls, the necessary conditions which the optimal controls must satisfy, and the computational aspects for computing the optimal controls are considered.